

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2005, 13:24:07 ; Search time 15065 Seconds

(without alignments)
12792.525 Million cell updates/sec

Title: US-09-931-375A-1

Perfect score: 5063
1 gccatgagagccgcgagtcgagc.....agggctcgagagacttgta 5063

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gest1: *
9: gb_gest2: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	2059	40.7	2838	3	AK090382 Mus muscu
2	1365.8	27.0	2752	3	AK081025 Mus muscu
3	932	18.4	1045	5	BM920216 AGENCOURT
4	919.6	18.2	1032	5	BM920216 AGENCOURT
5	914.8	18.1	1031	5	BM920216 AGENCOURT
6	877.8	17.3	1052	5	BM920216 AGENCOURT
7	836.4	16.5	888	5	BM920216 AGENCOURT
8	811	16.0	967	5	BM920216 AGENCOURT
9	807	15.9	870	5	BM920216 AGENCOURT
10	805	15.9	1001	5	BM920216 AGENCOURT
11	802.2	15.8	986	4	BM06571 AGENCOURT
12	785.2	15.5	932	4	BM06571 AGENCOURT
13	784.4	15.5	902	5	BM920216 AGENCOURT
14	783.4	15.5	855	5	BM920216 AGENCOURT
15	781.6	15.4	5400	9	AY407796 Homo sapi
16	779	15.4	911	5	BM920216 AGENCOURT
17	778.4	15.4	891	5	BM920216 AGENCOURT
18	776.8	15.3	866	5	BM920216 AGENCOURT
19	771.6	15.2	5400	9	AY407797 Pan trogl
20	766.4	15.1	880	7	CK769878 AGENCOURT
21	762.8	15.1	902	5	BM920216 AGENCOURT
22	758	15.0	886	6	CA489007 AGENCOURT
23	732	14.5	847	5	BM920216 AGENCOURT
24	723.6	14.3	5400	9	AY407798 Mus muscu

25	717.2	14.2	877	5	BM920216 AGENCOURT
26	711	14.0	876	5	BM920216 AGENCOURT
27	707.4	13.9	943	5	BM920216 AGENCOURT
28	704.4	13.9	706	5	BM920216 AGENCOURT
29	702.8	13.9	776	4	BM920216 AGENCOURT
30	702.6	13.9	1026	2	BM920216 AGENCOURT
31	699.4	13.8	834	1	AL525430
32	696.6	13.8	995	4	BM920216 AGENCOURT
33	688.2	13.6	956	5	BM920216 AGENCOURT
34	687.6	13.6	712	5	BM920216 AGENCOURT
35	684.4	13.5	821	1	AL525430
36	677.6	13.4	860	4	BM920216 AGENCOURT
37	673	13.3	1212	4	BM920216 AGENCOURT
38	671.8	13.3	921	4	BM920216 AGENCOURT
39	670.8	13.2	1076	2	BM920216 AGENCOURT
40	661.8	13.1	699	7	CF131821
41	652.6	12.9	854	5	BM920216 AGENCOURT
42	645.2	12.7	919	6	CA495285
43	636.8	12.6	850	5	BM920216 AGENCOURT
44	636.2	12.6	655	6	CA12718
45	635.6	12.6	938	2	BM920216 AGENCOURT

ALIGNMENTS

RESULT 1	AK090382	2838 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK090382				
DEFINITION	Mus musculus adult male intestinal mucosa cDNA, RIKEN full-length enriched library, clone:G630070K03 product:low density lipoprotein receptor-related protein 5, full insert sequence.				
ACCESSION	AK090382.1 GI:26105912				
VERSION	AK090382.1				
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subcloning of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata, K., Itoh, M., Atzawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunuma, T., Teshiro, H., Itoh, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	20530913				
PUBMED	11076861				
REFERENCE	4				
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409, 685-690 (2001)				
MEDLINE	5				
PUBMED	The FANTOM Consortium and the RIKEN Genome Exploration Research				

TITLE Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE (bases 1 to 2838)
AUTHORS Adachi,J., Aikawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,U., Kojima,Y., Kondo,S., Konno,H., Konda,M.,
Koyama,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Nunataki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku,Akashira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toyota,T., Yasunishi,A.,
Muzumatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,
URL: http://genome-gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome-gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
Source Location/Qualifiers
1..2838
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="PANTOM:DB:G630070K03"
/db_xref="taxon:10090"
/clone="G630070K03"
/sex="male"
/tissue_type="intestinal mucosa"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature
/note="Low density lipoprotein receptor-related protein 5
(MDJ:MG1:1278315, GI:NM_008513, evidence: BLASTN, 99%,
match=2144)
putative"
ORIGIN
Query Match 40.7%; Score 2059; DB 3; Length 2838;
Best Local Similarity 85.4%; Pred. No. 0;
Matches 2424; Conservative 0; Mismatches 375; Indels 40; Gaps 10;
QY 2230 TGGATGGGCAAGACCTTACTGCGGCGGACACTGGGACCAAGATGAGTGGCGCG 2289
DB 2 TGGATGGGCAAGACCTTACTGCGGCGGACACTGGGACCAAGATGAGTGGCGCG 61
QY 2290 CTGACGCGGACGTTCCGCGACAGTCTCGTGTGAGGAGACTTGACCAACCCGAGTGGT 2348
DB 62 CTGATGGGCGAGTTCCGCGACAGTCTGTGTGAGAGAGCTTGACCAACCCGAGTGGT 121
QY 2349 GGGCTTGATCCCAAGAGGCTACATCTGACCGAGTGGGCGGCAAGCGAGAT 2408
DB 122 GGGCTTGATCTCTAAAGGCTACATCTGACCGAGTGGGCGGCAAGGAT 181
QY 2409 CGTGGGCGGCTTTCATGACGCGGACCACTGATGACGCTGTGACCAAGT -GGGCGGG 2467
DB 182 TGTGGGCGGCTTTCATGATGAGGACCAATGTATGACATGTGTGACCAAGTGGGCGGG 241
QY 2468 CCAAGACCTTCACTGATGATGACGCTGACGAGCGCTTCTACTGACCGACCTGACACA 2527

DB 242 CCAAGACCTTCACTGATGATGATGACGCTGACGAGCTGATGACGAGCTGACGAGCA 301
QY 2528 ACATGATCGAGTGTCTCAACATGCTGGGTGAGAGCGGGTGTGATGGCCAGCATCTCC 2587
DB 302 ACATGATGAGTGTCTCAACATGCTGGGTGAGAGCGGGTGTGATGGCCAGCATCTCC 361
QY 2588 CGCACCCTGGTGTGAGCGAGTACAGCGATTAATATATCTACTGACAGACTGGAATCTGC 2647
DB 362 CTTACCCCTGGTGTGAGCGAGTACAGCGATTAATATATCTACTGACAGACTGGAATCTGC 421
QY 2648 ACAGATGAGCGGGCGGACCAAGATGACGCGGGGACCGGACCTCTACCGAGGCGCAC 2707
DB 422 ATAGATTTGAACGGGCGGACCAAGACAGTGGGCGGACCGGACCTCTACCGAGGCGCAC 481
QY 2708 TGAATCTGTGATGACATCTGTGTGCTCACTCTCCCGCAGAGATGGCTCAATGACT 2767
DB 482 TGAATCTGTGATGACATCTGTGTGCTCACTCTCCCGCAGAGATGGCTCAATGACT 541
QY 2768 GTATGCAACAACGCGGACGTTGGGCACTGTGCTTGGCATCCCGGCGGCGCACCGCT 2827
DB 542 GGTGCAACAACGCGGACGTTGGGCACTGTGCTTGGCATCCCGGCGGCGCACCGCT 601
QY 2828 GCGGCTGGCCCACTACAC -CCTGACCCCGGACCGGACCGGACCGGCGCACCG 2886
DB 602 GTGGCTGTGCTTCACTACACCGGCTGAGACCGGACCGGACCGGACCGGCGCACCG 653
QY 2887 ACCTTCTGTGCTTCACTACCGGACCGGACCGGACCGGACCGGACCGGCGCACCG 2946
DB 654 -----AAATTCATACCGGAGATGCCGATGACCAAGCTC 692
QY 2947 AGCCGGATCTCATCTCTCCCTGTGATGACTGAGGAACGTGAAGCATGACTATGAC 3006
DB 693 AGCCGGATCTCATCTCTCCCTGTGATGACTGAGGAACGTGAAGCATGACTATGAC 752
QY 3007 CCACTGCAAGATTCATCTCTGAGTGGGATGGGCGGCAACATCAACGCGCAAGGAC 3066
DB 753 CCGCTGCAAGATTCATCTCTGAGTGGGATGGGCGGCAACATCAACGCGCAAGGAC 812
QY 3067 GACGGGACCGACCCCTTGTGTTTGAACCTCTGTGACCGCAAGGCGCAACCGCAGCGAG 3126
DB 813 GACGGTACCGACCCCTTGTGTTTGAACCTCTGTGACCGCAAGGCGCAACCGCAGCGAG 872
QY 3127 CCGGACCGACCGGATGATGATCTACACCGGACGCTGTTGTGAGCTGGAGGCGCAC 3186
DB 873 CCGGACCGACCGGATGATGATCTACACCGGACGCTGTTGTGAGCTGGAGGCGCAC 932
QY 3187 AATACATCAACGTTCAACAGGCTGAGCGGGAACGATGGGGGTGTGCTGTGGGGAC 3246
DB 933 AATACATCAACGTTCAACAGGCTGAGCGGGAACGATGGGGGTGTGCTGTGGGGAC 992
QY 3247 CCGGACCGGACCGGATGATGATCTACACCGGACGCTGTTGTGAGCTGGAGGCGCAC 3306
DB 993 CCGGACCGGACCGGATGATGATCTACACCGGACGCTGTTGTGAGCTGGAGGCGCAC 1052
QY 3307 ATGAGAGACCGGAGCGCAAGATGGAACGCGGACCGCTGACCGGACCGGAGGCTC 3366
DB 1053 ATGAGAGACCGGAGCGCAAGATGGAACGCGGACCGCTGACCGGACCGGAGGCTC 1112
QY 3367 CTTTCAACACCGGCTCATCTCGGCTGTGGGCTGTGGGTGAGCAACACTGGGCAAG 3426
DB 1113 CTTTCAACACCGGCTCATCTCGGCTGTGGGCTGTGGGTGAGCAACACTGGGCAAG 1172
QY 3427 CTGTTCTGGGTGAGCGGACCTGAGCGCATTTGAGAGCTGTGACGTTGACGAGGCGCAAC 3486
DB 1173 CTGTTCTGGGTGAGCGGACCTGAGCGCATTTGAGAGCTGTGACGTTGACGAGGCGCAAC 1230
QY 3487 CGCTGACCGCTGAGAGCGCAACATGTGTGAGCTCTGGGCTGACATCTTGGCAAG 3546
DB 1231 CGCTGACCGCTGAGAGCGCAACATGTGTGAGCTCTGGGCTGACATCTTGGCAAG 1290
QY 3547 CATCTACTGATGACCGGCGGACGAGCATGATGACGCTGTGAGAGAGACCAAGCGGG 3606

Db 1291 CACCTCTACTGATCGACCGCCAGCAGCAGATGATCGAGCGCGTGGAGAAACCACTGGG 1350
 Qy 3607 GACAAAGCGACTCGCATCCAGGCGCGTGTGCGCCACCTCCTCCTGATCATCATGATGAG 3666
 Db 1351 GACAAAGCGACTAGGGTTGAGGCGCGTGTGCGCCACCTGATGAGGATCATGCGGT -GAG 1409
 Qy 3667 GAAATCAGCTGAGGAGTTCTCAGCCCAACCATGTCGCCGTGACAATGGTGGTGTCTCC 3726
 Db 1410 GAAATCAGCTGAGGAGTTCTCAGCCCAACCATGTCGCCGTGACAATGGTGGTGTCTCC 1469
 Qy 3727 CACATCTCTATTTGCAAGGCTGATGGGACACACACCGTGTCTATGCGCCAGTCCACCTCGT 3786
 Db 1470 CACATCTGTATGCGCAAGGCTGATGAAACACCGCTGTCTGTGCGCTGTCCACCTGGT 1529
 Qy 3787 CTCCTGCAAGACCTCTGATGAGAGCGCGCCACCTGTCTCCCGGACCAAGTTTGA 3846
 Db 1530 CTCCTGCAAGACCTGTGATGAGTGTGAGCTCTCTACTGTCTCCCTGATCAGTTTGA 1589
 Qy 3847 TGTGCGCACAGGGAGATCGACTGTATCCCGGGGCTGTGCGTGTGAGCGGCTTTCCCGAG 3906
 Db 1590 TGTACCACTGGTGAATCGACTGATCCCGGAGCTGTGAGCGGCTTCCCTCGAG 1649
 Qy 3907 TGCATGACCAAGAGCAGAGAGGCTGCGCGGTGTGCTGCGCGCCAGTTCCTCCCTGC 3966
 Db 1650 TGTGCTGACCAAGGCTGATGAAAGAGGCTGCCAGTGTGCTCGCGCTCTCAGTTCCCTGC 1709
 Qy 3967 GCGGGGCTCAGTGTGCGACTGCGCTGCGCTGCGAGCGGAGGAGCAGTGTCCAGAC 4026
 Db 1710 GCTCAGGCGCAGTGTGCGACTGCGGTGAGCTGCGAGCGGTGAGCGGCTGCGCAGGAT 1769
 Qy 4027 CGCTCAGCAGGCGGCACTGTGACCCCATCTGCTGCGCCCAACCACTTCCGGTGTGCGAGC 4086
 Db 1770 CGCTCTGATGAAGCTAACTGGATGCTGTCTGCTGCGCCCATCAGTTCCGGTGTGCGAGC 1829
 Qy 4087 GCGCAGTGTCTCATCAACAGCAGTGTGAGTCTCTTCCCGGCTGTATGAGCGTCTCC 4146
 Db 1830 GCGCAGTGTCTCATCAACAGCAGTGTGAGTCTCTTCCCGGCTGTATGAGCGTCTCC 1889
 Qy 4147 GACGAGCTCATGTGTGAATCACCAGCGCGCTCAGACGACAGCGCGGCCACAGCAGT 4206
 Db 1890 GATGAGCTCATGTGTGAATCAACAGCGCACCTCTGATGACATCCCAAGCCACAGCAGT 1949
 Qy 4207 GCCATCGGCGCGCTCATTTGGCATCATCTCTCTCTCTGTCATGGTGTGTCTATTTT 4266
 Db 1950 GCCATCGGCGCGCTCATTTGGTATCATCTCTCTCTCTGTCATGGTGTGTCTATTTT 2009
 Qy 4267 GTGTGCGCAGCGTGTGTGCGAGCGCTATGCGGGGCGCAACAGCGGCTTCCCGACAG 4326
 Db 2010 GTCTGCCAGCGTGTGTGCGAGCGCTACACAGGGGCGAGTGGGCGCTTCCCGACAG 2069
 Qy 4327 TATGTCAGGGGACCCCGCAGTGTGCGCTCTCAATTTATAGCCCGGGCGGTTCAGCAT 4386
 Db 2070 TATGTTGTGAGCGCCCTCATGTGCTCTCTCACTTCTAGCCCGGCGGTTCACAGC 2129
 Qy 4387 GGC -CCCTTTCACAGCATCGATGCGGAAAGTCCATGATGAGTCTCGTGAGCCTGATGGG 4445
 Db 2130 GGCTCCCTTCCAGGATCCCGTGAGCAGTGTGCTGATGAGTCTCATGAGCCTGGTGG 2189
 Qy 4446 GGGCGGGGCGGGGTG -CCCTGTACGACCGGAACCATGTCACAGGGGCGCTCGTCCAGCA 4504
 Db 2190 GGGCGGGGCGAGTGTGCGCCCTCTATGACCGGAATCATGTCAGTGGGCGCTCATCCAGCA 2249
 Qy 4505 GCTGTTCAGCAGCAGGCGCAGCTGTACCGCGCATCTGAAACCGCGCGCTTCCCGG 4564
 Db 2250 GCTGTTCAGCAGCAGGCGCAGCTATATCGCGCATCTGAAACCGCGCTTCCCGG 2309
 Qy 4565 CCAGGACCTCTCTGTACAACTGAGCATGTTCTACTTCTTCAAACTTCGCGCCACTG 4624
 Db 2310 CCAGGACCTCTCTCTACAACTGAGCATGTTTATTTCTTACGAGCTCCCGGCGCC 2369
 Qy 4625 CGAGACCGTACAGGCGCTTACATCTTCAGGAATGGCGCGCCCGCAGCAGCGCTGAGCA 4684
 Db 2370 CTAGACCATACAGGCGCTTACGTCATTCGAGGTATGGCACCCCGCAGCAGCGTGCAGCA 2429

Qy 4685 CCGAGCTGTGTGACAGCGACTACAGCGCAGCGCTGGAAGGCCAGCAAGTACTACCTGG 4744
 Db 2430 CAGATGTGTGTGACAGTGTGACTACAGCAGCAGTGTGTGAGAGCAGCAATACTACTGG 2489
 Qy 4745 ATTGAACTCGGACTCAGACCCCTTATCCACCCCAACCCACCGCCCAAGTACTGT 4804
 Db 2490 ACTTGAATTCGACTCAGACCCCTTATCCACCCCAACCCCAAGTACTGT 2549
 Qy 4805 CGGCGGAGCAGCTGCGCGCGCTTATCCACCGCAGCAGGAGTACTTCCATCTCTCC 4864
 Db 2550 CTGACAGAGCAGCTGCGCGCGCTTATCCACCGCAGCAGGAGTACTTCCATCTCTCC 2609
 Qy 4865 CGCCCTCTCTGCGCGCGCTTATCCACCGCAGCAGTCTATCTGCGCGCGCTTCTCTCT 4924
 Db 2610 CGCCCTCTCTGCGCGCGCTTATCCACCGCAGCAGTCTATCTGCGCGCGCTTCTCTCT 2667
 Qy 4925 CTGTGCGCTCTGAAATAGTATTAATATGAACAAAGAAAATAATATATATATATTA 4984
 Db 2668 CTGCTCTCTGAAATATTTTAAATATGAACAAAG -GAAAAATATATATATATATTA 2726
 Qy 4985 AAAATAATATATATTTGGGATTTTAAACATGAGAAATGTGAACTGTGATGGGTGGCA 5044
 Db 2727 AAAATAATATATATTTGGGATTTTAAACATGAGAAATGTGAGCGGTGAAGGGGTGGCA 2786
 Qy 5045 GGGCTGGGAGAACTTTGTA 5063
 Db 2787 GGGCTGGGAGAACTTTGTA 2805

RESULT 2

AK081025

LOCUS

DEFINITION

AK081025

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

AK081025 2752 bp mRNA linear HTC 03-APR-2004
 Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length
 enriched library, clone:B93069L06 product:low density lipoprotein
 receptor-related protein 5, full insert sequence.

AK081025
 HTC; CAP trapper.
 Mus musculus (house mouse)

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)

99279253
 10349636

2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)

20499374
 11042159

3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multipillar sequencer
 Genome Res. 10 (11), 1757-1771 (2000)

20530913
 11076861

4
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.


```

Db      881  GACTGTGACGNCATCTGCTG-CAAAACANTTCGGTGTGCCAGCGGCATTTGTCTCTC 939
Qy      4102 ATCAAAACAGCAGTGCAGTCTCTTCCCGACTGTATTCAGAGCGTTCGAGAGCTCATGTGT 4161
Db      940  ATCAAAAC-GCAGTGGCGGTCTCTTCCCGACTGTATTCAGAGCGTTCGAGAGCTCATGTGGG 998
Qy      4162 GAAATCACCAAGCGCCCTCAGACACAGCCCGGCCACAGCAGTG 4207
Db      999  GAAATCACAGCGCCCTCAGAACGAGCCCGGCCCAAAACAGTG 1044

RESULT 4
LOCUS   BX439521                1032 bp      mRNA      linear      EST 04-MAY-2004
DEFINITION BX439521 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE011YM02
5-PRIME, mRNA sequence.
ACCESSION BX439521
VERSION   BX439521.2 GI:47020909
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1032)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 15, 2003 this sequence version replaced gi:30787803.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 6245.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna/s=CS0DE011BG01QPI&c=6245.f.
Location/Qualifiers
1..1032
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE011YM02"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

FEATURES
source
Query Match 18.2%; Score 91.9.6; DB 5; Length 1032;
Best Local Similarity 99.4%; Pred. No. 5.6e-190;
Matches 930; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

Qy      1323 CTGGGTGGCCGAAACCTCTACTGACCGACACGCGGACCGCATCGAGGTGACGCG 1382
Db      1 CTGGGTGGCCGAAACCTCTACTGACCGACACGCGGACCGCATCGAGGTGACGCG 60
Qy      1383 CCTCAACGGACCTCCCGCAAGATCTCTGTCGAGGACCTGACGAGCCCGAGCCAT 1442
Db      61 CCTCAACGGACCTCCCGCAAGATCTCTGTCGAGGACCTGACGAGCCCGAGCCAT 120
Qy      1443 CGCACTGCACCCCGGTGATGGGCTCATGTCTGACGACCTGGGAGAGACCTTAAAT 1502
Db      121 CGCACTGCACCCCGGTGATGGGCTCATGTCTGACGACCTGGGAGAGAACCTTAAAT 180
Qy      1503 CGAGTGTGCCAACTTGGATGGCAGGACCGCGTGTGCTGCTCAATGCTCCCTCGGGTG 1562

```

```

Db      181  CGAGTGTGCCAACTTGGATGGGAGGAGCGCGTGTGCTGCTCAATGCTCCTCCCTCGGGTG 240
Qy      1563 GCCCAAACCGGCTCGGCTCTGGACCTGTCAGGAGGGGAGAGCTCTACTGGGGAGAGCGCAAAGAC 1622
Db      241  GCCCAAACCGGCTCGGCTCTGGACCTGTCAGGAGGGGAGAGCTCTACTGGGGAGAGCGCAAAGAC 300
Qy      1623 AGACAAGATCGAGTGCATGATGTTGATGGGACGAGAGGGGGAGCCCTCTCTGGAGGACAA 1682
Db      301  AGACAAGATCGAGTGCATGATGTTGATGGGACGAGAGGGGGAGCCCTCTCTGGAGGACAA 360
Qy      1683 GCTCCCGCACATTTTCGGGTTTCAGCTGCTGGGGGAGACTTCTACTTACTGAGACTGACTGGCA 1742
Db      361  GCTCCCGCACATTTTCGGGTTTCAGCTGCTGGGGGAGACTTCTACTTACTGAGACTGACTGGCA 420
Qy      1743 GCGCGCAGCATTCGAGCGGGTGCACAAGGTCAAGGCCAGCGGGGACGTCATCATTTGACCA 1802
Db      421  GCGCGCAGCATTCGAGCGGGTGCACAAGGTCAAGGCCAGCGGGGACGTCATCATTTGACCA 480
Qy      1803 GCTGCCCGACCTGATGGGGTCAAGGCTGTAAGTGTGGCCCAAGGTCGTCGGAACCAACCC 1862
Db      481  GCTGCCCGACCTGATGGGGTCAAGGCTGTAAGTGTGGCCCAAGGTCGTCGGAACCAACCC 540
Qy      1863 GTGTGCGGACAGGAAACGGGGGTGCAGCCACTGTGCTTCTTCAACCCCGACCGAACCCG 1922
Db      541  GTGTGCGGACAGGAAACGGGGGTGCAGCCACTGTGCTTCTTCAACCCCGACCGAACCCG 600
Qy      1923 GTGTGGCTGCCCATCGGCTTGGAGTGTGTCGAGTGCATGAAAGACCTGTCATCGTCCCTGA 1982
Db      601  GTGTGGCTGCCCATCGGCTTGGAGTGTGTCGAGTGCATGAAAGACCTGTCATCGTCCCTGA 660
Qy      1983 GGCCTTCTTGTGCTTCCACAGCAGCGCGCCATCCACAGGATCTCCCTCGAGACCAATAA 2042
Db      661  GGCCTTCTTGTGCTTCCACAGCAGCGCGCCATCCACAGGATCTCCCTCGAGACCAATAA 720
Qy      2043 CAACGAGCTGGCCATCCCGCTCAGCGGCGTCAAGGAGGCTCAGCCCTGGACTTTTGATGT 2102
Db      721  CAACGAGCTGGCCATCCCGCTCAGCGGCGTCAAGGAGGCTCAGCCCTGGACTTTTGATGT 780
Qy      2103 GTCCAAACACCATCTACTTGGACAGAGCTGAGCTGAGACCATGAGACCATGAGCCGCGCTTAT 2162
Db      781  GTCCAAACACCATCTACTTGGACAGAGCTGAGCTGAGACCATGAGACCATGAGCCGCGCTTAT 840
Qy      2163 GAACGGAGCTCGGTGGAGCAGCTGTTGGCTTTGGCTTACCTACCCCGAGGGCATGCG 2222
Db      841  GAACGGAGCTCGGTGGAGCAGCTGTTGGCTTTGGCTTACCTACCCCGAGGGCATGCG 900
Qy      2223 CGTTGACTGGATGGGCAAGAACCTCTACTGGGCGCA 2258
Db      901  CGTTGACTGGATGGGCAAA-AACTCTACTGGGCGCRM 935

RESULT 5
LOCUS   BX402084/c              1031 bp      mRNA      linear      EST 29-APR-2004
DEFINITION BX402084 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC011YB06 3-PRIME, mRNA sequence.
ACCESSION BX402084
VERSION   BX402084.2 GI:46874938
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1031)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 13, 2003 this sequence version replaced gi:30618689.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

```



```
QY 2557 CAGGAGCGGGTGGTGAATGTCGACAGCATCTCCGACACCGTTGCGTCTGACGAGTACAGC 2616
Db 421 CAGGAGCGGGTGGTGAATGTCGACAGCATCTCCGACACCGTTGCGTCTGACGAGTACAGC 480
QY 2617 GATTATATCTACTGACAGACTGGAACTGACAGCATTTGACGCGGCGGACAGACTAGC 2676
Db 481 GATTATATCTACTGACAGACTGGAACTGACAGCATTTGACGCGGCGGACAGACTAGC 540
QY 2677 GCGCGGAACCGCACCTCATCCAGGCGCACCTGGACTCTGATGAGACATCTGCTGTTG 2736
Db 541 GCGCGGAACCGCACCTCATCCAGGCGCACCTGGACTCTGATGAGACATCTGCTGTTG 600
QY 2737 CACTCTCTCCGCGGAGTGGCTCAATGACTGTATGACAAACAGGCGAGTGTGGGCGAG 2796
Db 601 CACTCTCTCCGCGGAGTGGCTCAATGACTGTATGACAAACAGGCGAGTGTGGGCGAG 660
QY 2797 CTGTGCTTGGCATCCCGGGGCGCACCGCTGCGGCTGCGCTCACTACACCTTGGAC 2856
Db 661 CTGTGCTTGGCATCCCGGGGCGCACCGCTGCGGCTGCGCTCACTACACCTTGGAC 720
QY 2857 CCCAGCAGCGCAACTGCGAGCCCGCCACACCTTCTGCTGTTGTCAGCCAGAAATCTGCC 2916
Db 721 CCCAGCAGCGCAACTGCGAGCCCGCCACACCTTCTGCTGTTGTCAGCCAGAAATCTGCC 780
QY 2917 ATCAGTCGAGTATCCCGGAGCAGCAGCAGCCCGGATCTCATCTCTGCTGCTGATGG 2975
Db 781 ATCAGTCGAGTATCCCGGAGCAGCAGCAGCCCGGATCTCATCTCTGCTGCTGATGG 840
QY 2976 ACTGAGGAGCTCAAGCAGTATGATGACC 3007
Db 841 GACTGAGGAGCTCAAGCAGTATGATGACC 872

RESULT 8
BX428434 967 bp mRNA linear EST 05-MAY-2004
LOCUS BX428434 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE011YM02
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION BX428434
VERSION BX428434.1 GI:31017882
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 967)
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 6245.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?e=CS0AQ006ZH04_A0512_1&c=6245.f.
Location/Qualifiers
1. 967
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE011YM02"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector."

FEATURES
source
```

```
ORIGIN
Library was not normalized."
Query Match 16.0%; Score 811; DB 5; Length 967;
Best Local Similarity 90.3%; Pred. No. 3.1e-166;
Matches 859; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1874 GGAACGGGGGTGACGACCTGCTGCTTCTTTCACACCCACGCAACCCGGTGTGCTGCTGCC 1933
Db 6 GGTACGGGGGTGACGACCTGCTGCTTCTTTCACACCCACGCAACCCGGTGTGCTGCTGCC 65
QY 1934 CCATCGGCTCGAGCTGCTGAGTGAATGAAGACCTGCTGCTGCTGAGGCTTCTTTGG 1993
Db 66 CCATCGGCTCGAGCTGCTGAGTGAATGAAGACCTGCTGCTGCTGAGGCTTCTTTGG 125
QY 1994 TCTTCAACGACGAGACCGCCCATCCACAGGATCTCCCTCGAGACCAATTAACACGAGCTGG 2053
Db 126 TCTTCAACGACGAGACCGCCCATCCACAGGATCTCCCTCGAGACCAATTAACACGAGCTGG 185
QY 2054 CCATCCCGCTCACGGGGGTCAAGGAGGCTCAGCCCTGGACTTTTGTATGTGTCCAAACACC 2113
Db 186 CCATCCCGCTCACGGGGGTCAAGGAGGCTCAGCCCTGGACTTTTGTATGTGTCCAAACACC 245
QY 2114 ACATCTTACTGACAGAGCTCAGCCCTGAAGACCATCAGCCGCGCTTTCATGAACGGGAGCT 2173
Db 246 ACATCTTACTGACAGAGCTCAGCCCTGAAGACCATCAGCCGCGCTTTCATGAACGGGAGCT 305
QY 2174 CGGTGGAGCAGCTGGTGGAGTTTGGCTTGTACTACCCCGAGGGCATGCGCTTGTATGGA 2233
Db 306 CGGTGGAGCAGCTGGTGGAGTTTGGCTTGTACTACCCCGAGGGCATGCGCTTGTATGGA 365
QY 2234 TGGGCAAGAACTCTTACTTGGGCGGACACTGGGACCAACAGAAATCGAAGTGGCGCGCTGG 2293
Db 366 TGGGCAAGAACTCTTACTTGGGCGGACACTGGGACCAACAGAAATCGAAGTGGCGCGCTGG 425
QY 2294 ACGGCGAGTTTCCGGCAAGTCTCTGTGTGGAGGGACTTTGGACAACCCCGAGGTCGCTGGCCC 2353
Db 426 ACGGCGAGTTTCCGGCAAGTCTCTGTGTGGAGGGACTTTGGACAACCCCGAGGTCGCTGGCCC 485
QY 2354 TGGATCCCAACGAGGCTACATCTACTTGGACCGAGTGGGGCGGACGAGGATCGTGC 2413
Db 486 TGGATCCCAACGAGGCTACATCTACTTGGACCGAGTGGGGCGGACGAGGATCGTGC 545
QY 2414 GGGCTTTCATGAGCGGACCAACTGTCATGAGCTGTGGACAAGTGGCGCGGCGCAACG 2473
Db 546 GGGCTTTCATGAGCGGACCAACTGTCATGAGCTGTGGACAAGTGGCGCGGCGCAACG 605
QY 2474 ACCTCACCATTTGACTACGCTGACCGGCTTCTACTGGACCGGACCTTGGACCAACATGA 2533
Db 606 ACCTCACCATTTGACTACGCTGACCGGCTTCTACTGGACCGGACCTTGGACCAACATGA 665
QY 2534 TCGAGTCTGTCACCATGCTGGTTCAGGACGGGCTGCTGATTTGCCGACGATCTCCCGAC 2593
Db 666 TCGAGTCTGTCACCATGCTGGTTCAGGACGGGCTGCTGATTTGCCGACGATCTCCCGAC 725
QY 2594 CGTTTCGGTCTGACGCGAGTACAGCGATATATCTACTGACAGACTTGAATCTGACACGA 2653
Db 726 CGTTTCGGTCTGACGCGAGTACAGCGATATATTTCTACTGACAGACTTGAATCTGACACGA 785
QY 2654 TTGAGCGGGCGGACCAAGACTAGCGGCGGGAACCGCACCCCTCATCCAGGCGCACCTGGACT 2713
Db 786 TTGAGCGGGCGGACCAAGACTAGCGGCGGGAACCGCACCTCATCCAGGNNCCCTGNACT 845
QY 2714 TCGTGTATGGAATCCTGCTGTTTCCATCTCTCCGCGAGGATGGGCTCAATGACTGTATGC 2773
Db 846 NCGTGTATGGAATCCTGCTGTTTCCCTCTCTNNCCGAGNAGNCCTCATGACTGTGTCACACC 905
QY 2774 ACAAACAGGCGAGTGTGGGCGAGCTGTGCTTGGCATCCCGGGGGGCCACC 2824
Db 906 TCGGCAAGTGTGGAGTTGGGCTTTCCTTTTCGCGGAGACCCGCTGCGNTTAGC 956

RESULT 9
```

```

BQ687102/c
LOCUS BQ687102 870 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT_8046500 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6209180
5', mRNA sequence.
ACCESSION BQ687102
VERSION BQ687102.1 GI:21812406
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 870)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2368 row: a column: 21
High quality sequence stop: 680.
FEATURES
source
1..870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6209180"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 110"
/notes="Organ: Pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
ORIGIN
Query Match 15.9%; Score 807; DB 5; Length 870;
Best Local Similarity 98.7%; Pred.No. 2.3e-165;
Matches 824; Conservative 0; Mismatches 10; Indels 1; Gaps 1;
3396 GGCCTGGTGGTGGACACACACT-GGGCAAGCTGTTCTGGTGGACGGGACCTGAAGC 3454
835 GCCCGTGGTGGTGGACACACACTGGGGCAAGCTGTTCTGGGGACCGCGGACCTGAAGC 776
3455 GCATTGAGAGCTGTGACCTGTACAGGGGCCAACCGCCTGACCTTGGAGGACGCCAATCG 3514
775 GCATTGAGAGCTGTGACCTGTACAGGGGCCAACCGCCTTGAACCTGGAGGACGCCAATCG 716
3515 TGCAGCCTCTGGGCTGACCATCTCTGGGAGCATCTCTACTGATCGACCGCCAGCAGC 3574
715 TGCAGCCTCTGGGCTGACCATCTCTGGGAGCATCTCTACTGATCGACCGCCAGCAGC 656
3575 AGATGATCGAGCGTGTGGAGAGACACACCGGGGACACAGCGACTCGCATCCAGGGCCGCTG 3634
655 AGATGATCGAGCGTGTGGAGAGACACACCGGGGACACAGCGACTCGCATCCAGGGCCGCTG 596
3635 TCGCCCACTCTCATTGGCATCCATGAGTGAGGAGTGCACCTGGAGGAGTTCTCAGCCC 3694
595 TCGCCCACTCTCATTGGCATCCATGAGTGAGGAGTGCACCTGGAGGAGTTCTCAGCCC 536
3695 ACCCATGCGCCGAGCAATGGTGGCTGCTCCACATCTCTATTGCCAAGGGTGATGGGA 3754
535 ACCCATGCGCCGAGCAATGGTGGCTGCTCCACATCTCTATTGCCAAGGGTGATGGGA 476
```

```

3755 CACACGGTGTCTATGCCACCTCCAGTCTCTCTCCAGAACCTGCTGACCTGTGGAG 3814
475 CACACGGTGTCTATGCCACCTCCAGTCTCTCTCCAGAACCTGCTGACCTGTGGAG 416
3815 AGCCGCCACCTGCTCCCGGAGACCAAGTTTGCATGTGCCACAGGGGAGATCGACTATCC 3874
415 AGCCGCCACCTGCTCCCGGAGACCAAGTTTGCATGTGCCACAGGGGAGATCGACTATCC 356
3875 CCGGGGCTTGGCGCTGTGACCGCTTCCGAGTGGATGACACAGGAGGAGGGCT 3934
355 CCGGGGCTTGGCGCTGTGACCGCTTCCGAGTGGATGACACAGGAGGAGGGCT 296
3935 GCCCGTGTGCTCCCGGCCCGAGTTCCCTGCGCGGGGTCAGTGTGTGGACCTTCCGCC 3994
295 GCCCGTGTGCTCCCGGCCCGAGTTCCCTGCGCGGGGTCAGTGTGTGGACCTTCCGCC 236
3995 TGGCTTGGCAGCGGCGGAGGAGACTGTTCAGGACCGCTCAGACGCGGAGCTGTGACGCCA 4054
235 TGGCTTGGCAGCGGCGGAGGAGACTGTTCAGGACCGCTCAGACGCGGAGCTGTGACGCCA 176
4055 TCTGCTGCGCAACCACTTCCTGGTGGCGGCGGAGTGTCTCTCATCAACACAGCT 4114
175 TCTGCTGCGCAACCACTTCCTGGTGGCGGCGGAGTGTCTCTCATCAACACAGCT 116
4115 GCGACTCTTCCCGGAGTGTATCGACGGTCCGACGAGCTCATGTGTGAATCACAAGC 4174
115 GCGACTCTTCCCGGAGTGTATCGACGGTCCGACGAGCTCATGTGTGAATCACAAGC 56
4175 GCGCTTCCGAGCAGCGCGGCCCGCCACAGCAGTCCATCGGCGGCTCATTTGGCAT 4229
55 GCGCTTCCGAGCAGCGCGGCCCGCCACAGCAGTCCATCGGCGGCTCATTTGGCAT 1
RESULT 10
BX402085 1001 bp mRNA linear EST 29-APR-2004
LOCUS BX402085 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DC011YB06 5-PRIME, mRNA sequence.
ACCESSION BX402085
VERSION BX402085.2 GI:46876774
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 1001)
AUTHORS Li W.B., Gruber C., Jesses, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 13, 2003 this sequence version replaced gi:30622712.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6245.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS1AC003ZF04QPI&c=6245.f.
Location/Qualifiers
1..1001
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC011YB06"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
```


ORIGIN sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 15.9%; Score 805; DB 5; Length 1001;
Best Local Similarity 96.3%; Pred. No. 6.5e-165;
Matches 873; Conservative 5; Mismatches 21; Indels 8; Gaps 5;

QY 2978 TGAGGAACGTCAAGCCATCGACTATGACCCCACTGGACAAGTTCTACTCTGCTGGTGGATG 3037
DB 4 TGAGGAACGTCAAGCCATCGACTATGACCCCACTGGACAAGTTCTACTCTGCTGGTGGATG 63

QY 3038 GCGCCGACAACTCAAGGAGCGCAAGGAGCGGACCGACCGCTTTGTTTGACCTCTC 3097
DB 64 GCGCCGACAACTCAAGGAGCGCAAGGAGCGGACCGACCGCTTTGTTTGACCTCTC 123

QY 3098 TGAGCAAGGCGCAAAACCCAGACGCGCCCGCCAGCTCAGCATCGACATCTACAGCC 3157
DB 124 TGAGCAAGGCGCAAAACCCAGACGCGCCCGCCAGCTCAGCATCGACATCTACAGCC 183

QY 3158 GGACACTGTTCTGGACGTGCGAGGCGCACCAATACCATCAAGTCCACAGGCTGAGCGGG 3217
DB 184 GGACACTGTTCTGGACGTGCGAGGCGCACCAATACCATCAAGTCCACAGGCTGAGCGGG 243

QY 3218 AAGCCATGCGGCTGCTGCTGGGAGCGCGGACCAAGCCAGGCGCATCGTGTCAAGC 3277
DB 244 AAGCCATGCGGCTGCTGCTGGGAGCGCGGACCAAGCCAGGCGCATCGTGTCAAGC 303

QY 3278 CGGAGCGAGGTGACCTGCTACCAACATCGAGCGCGGCGGACCGAAGTCCAGACCG 3337
DB 304 CGGA--SAGGCTGCTGCTGCTACCAACATCGAGCGCGGCGGACCGAAGTCCAGACCG 361

QY 3338 CAGCCCTGGAGCGGACCGAGCGCGAGGCTCTCTTCCACACCGGCTCATCGCCCTGTGG 3397
DB 362 CA-CCCTGGAGCGGACCGAGCGGAGGCTCTCTTCCACACCGGCTCATCGCCCTGTGG 420

QY 3398 CCTGCTGGTGAGCAACAACACTGGGCAAGCTGTTCTGGGTGAGCGCGCACTGGAAGCGCA 3457
DB 421 CCTGCTGGTGAGCAACAACACTGGGCAAGCTGTTCTGGGTGAGCGCGCACTGGAAGCGCA 480

QY 3458 TTGAGAGCTGACCTGTCAGGCGGCAACCGCTGACCTGGAGGACCGCAACATCGTGC 3517
DB 481 TTGAGAGCTGACCTGTCAGGCGGCAACCGCTGACCTGGAGGACCGCAACATCGTGC 540

QY 3518 AGCCTCTGGGCTGACCATCTTGGCAAGCATCTCTACTGATCGACCGCAGCAGCAGA 3577
DB 541 AGCCTCTGGGCTGACCATCTTGGCAAGCATCTCTACTGATCGACCGCAGCAGCAGA 600

QY 3578 TGATCGAGCGTGTGAGGAAGACCAACCGGCGGACCAAGCGGACCTCGCATCCAGGCGCGTGTGC 3637
DB 601 TGATCGAGCGTGTGAGGAAGACCAACCGGCGGACCAAGCGGACCTCGCATCCAGGCGCGTGTGC 660

QY 3638 CCCACCTCACTGGCATCCATGAGTGGAGGAGTCAAGTCTGAGGAGGTTCTCAGCCAC 3697
DB 661 CCCACCTCACTGGCATCCATGAGTGGAGGAGTCAAGTCTGAGGAGGTTCTCAGCCAC 720

QY 3698 CATGTGCCGTGACAAATGTTGCTCTCCACATCTGATTATGCCAAGGCTGATGSGAC 3757
DB 721 CATGTGCCGTGACAAATGTTGCTCTCCACATCTGATTATGCCAAGGCTGATGSGAC 779

QY 3758 CACGCTGCTCATGCTCCCGGACGAGTTGATGTCACAGAGGAGTTCGACTGTATCCCG 3817
DB 780 CACGCTGCTCATGCTCCCGGACGAGTTGATGTCACAGAGGAGTTCGACTGTATCCCG 838

QY 3818 GCGCCACCTGCTCCCGGACGAGTTGATGTCACAGAGGAGTTCGACTGTATCCCG 3877
DB 839 GCGCCACCTGCTCCCGGACGAGTTTTCATGT---BCAMARGGAGATCGATGTTTCCCGGG 895

QY 3878 GGGCTG 3884
DB 896 GCTGCT 902

RESULT 11

BM561582 986 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT 6567315 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5739763
DEFINITION 5', mRNA sequence.
ACCESSION BM561582
VERSION BM561582.1 GI:18806953
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 986)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12753 row: j column: 20
High quality sequence stop: 680.

FEATURES
source
1..986
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5739763"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_88"
/note="Organ: Small intestine; Vector: pCMV-SPORT6;
Site 1: NotI; Site 2: SalI; Cloned unidirectionally;
oligo-dT primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 15.8%; Score 802.2; DB 4; Length 986;
Best Local Similarity 98.6%; Pred. No. 2.7e-164;
Matches 831; Conservative 0; Mismatches 8; Indels 4; Gaps 2;

QY 1778 CCAGCCGGAGCTCATATTGACAGTGCCCGACCTGATGGGCTCAAGCTGTGAATG 1837
DB 1 CCAGCCGGAGCTCATATTGACAGTGCCCGACCTGATGGGCTCAAGCTGTGAATG 60

QY 1838 TGGCCAAAGTCTGTCGGACCAACCCGTGTGCGGACAGGAACGGGGGTGCAGCCACCTGT 1897
DB 61 TGGCCAAAGTCTGTCGGACCAACCCGTGTGCGGACAGGAACGGGGGTGCAGCCACCTGT 120

QY 1898 GCTTCTTTACACCCACGCAACCCGCTGTGCTGCCCATCGGCTGTGAGTGTCTGAGTG 1957
DB 121 GCTTCTTTACACCCACGCAACCCGCTGTGCTGCCCATCGGCTGTGAGTGTCTGAGTG 180

QY 1958 ACATGAAGACCTGTCATGCTGCTGAGGCTTTCTTGGTCTTCCACGACGAGCCGCAATCC 2017
DB 181 ACATGAAGACCTGTCATGCTGCTGAGGCTTTCTTGGTCTTCCACGACGAGCCGCAATCC 240

QY 2018 ACAGGATCTCCCTCGAGNACCAATACACGACGTTGGCCATCCCGCTCAGCGGCTCAAGG 2077
DB 241 ACAGGATCTCCCTCGAGNACCAATACACGACGTTGGCCATCCCGCTCAGCGGCTCAAGG 300

QY 2078 AGGCTCTAGCCCTGGACTTTGATGTGTCCAAACACCATCTACTGGACACAGCTCAGCC 2137
DB 301 AGGCTCTAGCCCTGGACTTTGATGTGTCCAAACACCATCTACTGGACACAGCTCAGCC 360

QY 2138 TGAAGACCATCAGCCGCGCTTTCATGAACGGGAGCTCGGTGGAGCACTGTGTGGAGTTTG 2197
DB 361 TGAAGACCATCAGCCGCGCTTTCATGAACGGGAGCTCGGTGGAGCACTGTGTGGAGTTTG 420

QY 2198 GCCTTACTACCCGAGGCGATGGCCCTTACTGATGGCAAGAACTCTACTGGCGG 2257
Db 421 GCCTTACTACCCGAGGCGATGGCCCTTACTGATGGCAAGAACTCTACTGGCGG 480
QY 2258 ACACCTGGGACCAACAGAAATCGAAGTGGCGCGCTGGACGGGCGAGTTCGGCAAGTCCCTCG 2317
Db 481 ACACCTGGGACCAACAGAAATCGAAGTGGCGCGCTGGACGGGCGAGTTCGGCAAGTCCCTCG 540
QY 2318 TGTGAGGAGCTTGACAAACCGAGGTCGTGGCCCTGGATCCCAACGAAGGCTACATCT 2377
Db 541 TGTGAGGAGCTTGACAAACCGAGGTCGTGGCCCTGGATCCCAACGAAGGCTACATCT 600
QY 2378 ACTGACACGAGTGGGGCGGCAAGCGAGGATCGTGGGGCGCTTCATGACGGGACCAACT 2437
Db 601 ACTGACACGAGTGGGGCGGCAAGCGAGGATCGTGGGGCGCTTCATGACGGGACCAACT 660
QY 2438 GCATGACGCTGGTGACAAAGTGGGCGCGGCAACGAGCTCACCAATTGACTACGCTGACC 2497
Db 661 GCATGACGCTGGTGACAAAGTGGGCGCGGCAACGAGCTCACCAATTGACTACGCTGACC 720
QY 2498 AGCGCTCTACTGACCGAGCTTGACAAACATGATCGAGTGGTCCAAACATGCTGGG 2555
Db 721 AGCGCTCTACTGACCGAGCTTGACAAACATGATCGAGTGGTCCAAACATGCTGGG 780
QY 2556 TCAGGAGGGTGGGATGTCGAGATCTCCGCA--CCGTTGCGTCTGACGCGATAC 2613
Db 781 CAGGAGGGTGGGATGTCGAGATCTCCGCAACCGTTCGGGCGCTGACGCGATAC 840
QY 2614 AGC 2616
Db 841 AGC 843

RESULT 12

BM006571/c
LOCUS
DEFINITION 603615546F1 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:5421467 5',
mRNA sequence.

ACCESSION BM006571

VERSION BM006571.1

KEYWORDS EST.

SOURCE GR:16520925

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 932)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCW1877 row: 1 column: 12

High quality sequence stop: 757.

Location/Qualifiers

1..932

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5421467"

/tissue_type="ductal carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_110"

/note="Organ: pancreas; Vector: pOTB7; Site:1: XhoI;

Site 2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match Similarity 15.5%; Score 785.2; DB 4; Length 932;
Best Local Similarity 96.5%; Pred. No. 1.4e-160;
Matches 867; Conservative 0; Mismatches 23; Indels 8; Gaps 6;
QY 3337 GCAGCCCTGGAGCGGACCGAGCGAGGTCTCTTCAACCGGCTCATCGCCCTGTG 3396
Db 895 GCGAGCTGGAGCGGACGAGCGGGTCCGTGGTCAACCGG-CTCAACGCGCCTGTG 837
QY 3397 GCGCTGGTGGTGGACAAACACACTGGGCAAGCT-GTTCGGGTGGACGCGGACCTGAAAGCG 3455
Db 836 G-CCTGGTGGTGGCAACACACTGGGCAAGCTGTCTTGGGGGACGCGGACCTGAAAGCG 778
QY 3456 CATTGAGAGCTGTGACCTGTTCAGGGGCGCAACCGGCTGACCTGGAGGAGC-CCACATCG 3514
Db 777 CATTGAGA--CAGGACCTGTTCAGGGGCGCAACCGGCTGACCTGGAGGAGCGCCCAACATCG 720
QY 3515 TGCA--GCCTCTGGGCGCTGACCATCTCTTGGCAAGCATCTCTACTGGATCGACCGCCAGCA 3572
Db 719 TGCAAGCTCTGGGCGCTGACCATCTCTTGGCAAGCATCTCTACTGGATCGACCGCCAGCA 660
QY 3573 GCAGATGATCGAGCGTGTGGAGAGAGCAACCGGGGAGCAAGCGGACTCGCATCCAGGGCGG 3632
Db 659 GCAGATGATCGAGCGTGTGGAGAGAGCAACCGGGGAGCAAGCGGACTCGCATCCAGGGCGG 600
QY 3633 TGTGCGCCACCTCACTGGCATCCATGAGTGGAGGAGTCAAGCTGGAGGAGTTCACG 3692
Db 599 TTTGCGCCACCTCACTGGCATCCATGAGTGGAGGAGTCAAGCTGGAGGAGTTCACG 540
QY 3693 CCACCCATGTGCCCGTGCATATGCTGTCTCCACATCTGTATTGCAAGGAGTGTATG 3752
Db 539 CCACCCATGTGCCCGTGCATATGCTGTCTCCACATCTGTATTGCAAGGAGTGTATG 480
QY 3753 GACACACGCGTGTCTATGCCCGAGTCCACCTGCTGCTCTGTCAGAAACCTGTGACTGTGG 3812
Db 479 GACACACGCGTGTCTATGCCCGAGTCCACCTGCTGCTCTGTCAGAAACCTGTGACTGTGG 420
QY 3813 AGAGCGCCCGACCTGTCTCCCGGACAGTTCGATGTGCCACAGGAGAGATCGACTGAT 3872
Db 419 AGAGCGCCCGACCTGTCTCCCGGACAGTTCGATGTGCCACAGGAGAGATCGACTGAT 360
QY 3873 CCGCGGGCGCTGGCGCTGTGACGCGCTTTCCCGAGTGGGATGACACAGAGCGAGGAGGG 3932
Db 359 CCGCGGGCGCTGGCGCTGTGACGCGCTTTCCCGAGTGGGATGACACAGAGCGAGGAGGG 300
QY 3933 CTGCGCGTGTGCTCCCGCGCCAGTTCCTCTGCGCGGGGTCACTGTGTGGAGCTCGG 3992
Db 299 CTGCGCGTGTGCTCCCGCGCCAGTTCCTCTGCGCGGGGTCACTGTGTGGAGCTCGG 240
QY 3993 CCGTGGCTGGAGCGGCGGAGCGAGTCTCAGGACCGCTCAGACGAGGCGGAGTGTGAGCG 4052
Db 239 CCGTGGCTGGAGCGGCGGAGCGAGTCTCAGGACCGCTCAGACGAGGCGGAGTGTGAGCG 180
QY 4053 CATCTGCTGCCCAACCAAGTTCGCGGTGTGGAGGCGGCGAGTGTCTCTCATCAACAGCA 4112
Db 179 CATCTGCTGCCCAACCAAGTTCGCGGTGTGGAGGCGGCGAGTGTCTCTCATCAACAGCA 120
QY 4113 GTGCGACTCTCTTCCCGAGCTGTATCGAGCGGTCCGAGAGCTCATGTGTGAATCACCAG 4172
Db 119 GTGCGACTCTCTTCCCGAGCTGTATCGAGCGGTCCGAGAGCTCATGTGTGAATCACCAG 60
QY 4173 GCGCGCTCAGACAGACCGCGGCCACAGCAGTGCATCGGGCCCGCTCATTTGGCATC 4230
Db 59 GCGCGCTCAGACAGACCGCGGCCACAGCAGTGCATCGGGCCCGCTCATTTGGCATC 2

RESULT 13

```
BUI95422
LOCUS       BUI95422       902 bp      mRNA      linear      EST 04-SEP-2002
DEFINITION  AGENCOURT 7591926 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6171960
5' mRNA sequence.
ACCESSION   BUI95422
VERSION     BUI95422.1 GI:22709406
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 902)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13541 row: c column: 01
            High quality sequence stop: 638.
FEATURES             Location/Qualifiers
     source           1..902
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6171960"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_67"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
                     Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.75 kb. Library constructed by Life
                     Technologies."
ORIGIN
Query Match      15.5%; Score 784.4; DB 5; Length 902;
Best Local Similarity 95.6%; Pred. No. 2e-160;
Matches 862; Conservative 0; Mismatches 31; Indels 9; Gaps 5;

QY  3520 CCTCTGGCCCTGACCATCTTGGCAGCATCTCTACTGATCGACCGCAGCAGATG 3579
DB  1 CCACGCGTCCGACCATCTTGGCAGCATCTCTACTGATCGACCGCAGCAGATG 60

QY  3580 ATCGAGCGGTGGAGAAGACACCGGGACAGCGGACTCGCATCCAGCGCGGTGCGCC 3639
DB  61 ATCGAGCGGTGGAGAAGACACCGGGACAGCGGACTCGCATCCAGCGCGGTGCGCC 120

QY  3640 CACCTCACTGGCATCCATGCGATGGAGGAAGTCAGCTGGAGGAGTTCTCAGCCACACCA 3699
DB  121 CACCTCACTGGCATCCATGCGATGGAGGAAGTCAGCTGGAGGAGTTCTCAGCCACACCA 180

QY  3700 TGTGCCCGGTGACATGGTGGCTGCTCCACATCTGTATTGCAAGGATGGAGCACCA 3759
DB  181 TGTGCCCGGTGACATGGTGGCTGCTCCACATCTGTATTGCAAGGATGGAGCACCA 240

QY  3760 CGGTGCTCATGCCAGTCCACTCGTGTCTCTGCAGAACCTGCTGACCTGTGGAGAGCGG 3819
DB  241 CGGTGCTCATGCCAGTCCACTCGTGTCTCTGCAGAACCTGCTGACCTGTGGAGAGCGG 300

QY  3820 CCCACTGCTCCCGGACCAAGTTTCATGTGTCACAGGGGAGATCGACTGATCCCGGG 3879
DB  301 CCCACTGCTCCCGGACCAAGTTTCATGTGTCACAGGGGAGATCGACTGATCCCGGG 360

QY  3880 GCCTGGCGCTGTGACGGCTTTCCCGAGTGCAGTACCGAGGACGAGGAGGCTGCCCC 3939
DB  361 GCCTGGCGCTGTGACGGCTTTCCCGAGTGCAGTACCGAGGACGAGGAGGCTGCCCC 420
```

```
QY  3940 GTGTGCTCCGCGGCCAGTTCCTCGCGCGGGGTGAGTGTGTGAGACTGCGCCTGGCC 3999
DB  421 GTGTGCTCCGCGGCCAGTTCCTCGCGCGGGGTGAGTGTGTGAGACTGCGCCTGGCC 480

QY  4000 TCGACGCGGAGGAGACTGTCAAGACCGCTCAGACGAGGCGGACTGTGAGCCCATGTC 4059
DB  481 TCGACGCGGAGGAGACTGTCAAGACCGCTCAGACGAGGCGGACTGTGAGCCCATGTC 540

QY  4060 CTGCCCAACCACTTCGGGTGCGAGCGGCGCAGTGTCTCTCATCAACAGCAGTGGCAG 4119
DB  541 CTGCCCAACCACTTCGGGTGCGAGCGGCGCAGTGTCTCTCATCAACAGCAGTGGCAG 600

QY  4120 TCCTTCCCGCACTGTATCGACGGCTCCGACGAGCTCATGTGTGAATCACCAGCGGCC 4179
DB  601 TCCTTCCCGCACTGTATCGACGGCTCCGACGAGCTCATGTGTGAATCACCAGCGGCC 660

QY  4180 TCAGACGACAGCCCGGCCACAGCAGTGCCTATCGGGCCGCTCATTTGGCATCATCTCTCT 4239
DB  661 TCAGACGACAGCCCGGCCACAGCAGTGCCTATCGGGCCGCTCATTTGGCATCATCTCTCT 720

QY  4240 CTCCTCGTCATGGGTGGTCTATTTTGTGTGCCAGCG-CGTGGTGTGCCAGCGTATGCG 4298
DB  721 CTCCTCGTCATGGGTGGTCTATTTTGGGTGCCCGCGCGGGGTGTGCCAACGCTATGCG 780

QY  4299 GGGGG--CCAAACGGGCGCTTCCCGCACGAGTATGTACAGCGGACCCCGCAC---GTGCC 4353
DB  781 GGGGGCCAAACGGGCGCTTCCCGCACGAGTATGTACAGCGGACCCCGCACGTGGGCC 840

QY  4354 CTCGAATTCATA-GCCCCGGCGGTTCCACAGCAT--GGCCCTTCACAGGATCCATGCG 4410
DB  841 CTCGAATTCATAACCCCGCGGTTCCACAGCATTTCCAGCATTTGGCCCTTCCAAAGCATTCATG 900

QY  4411 GG 4412
DB  901 GG 902

RESULT 14
BQ228206
LOCUS      BQ228206      855 bp      mRNA      linear      EST 02-MAY-2002
DEFINITION  AGENCOURT 7593499 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6020654
5' mRNA sequence.
ACCESSION   BQ228206
VERSION     BQ228206.1 GI:20409606
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 855)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13225 row: b column: 15
            High quality sequence stop: 608.
FEATURES             Location/Qualifiers
     source           1..855
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6020654"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_70"
```

/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.1 kb. Library constructed by Life Technologies."

ORIGIN

```
Query Match      15.5%; Score 783.4; DB 5; Length 855;
Best Local Similarity 97.9%; Pred. No. 3.3e-160;
Matches 836; Conservative 0; Mismatches 12; Indels 6; Gaps 4;

QY 3070 GGGACCCAGCCCTTGTGTTGACCTCTCTGAGCCAAAGCCAAACCCAGAGCAGGCC 3129
DB 1 GGGACCCAGCCCTTGTGTTGACCTCTCTGAGCCAAAGCCAAACCCAGAGCAGGCC 60

QY 3130 CACGACCTCAGCATCGACATCTACAGCCGGACACTGTTCTGAGCGTGGAGGCCACCAAT 3189
DB 61 CACGACCTCAGCATCGACATCTACAGCCGGACACTGTTCTGAGCGTGGAGGCCACCAAT 120

QY 3190 ACCATCAAGCTCCAGAGCTGAGCGGGGAGCCATGGGGGTGGTCTCGTGGGACCGC 3249
DB 121 ACCATCAAGCTCCAGAGCTGAGCGGGGAGCCATGGGGGTGGTCTCGTGGGACCGC 180

QY 3250 GACAGCCAGGGCCATCGTCTCAACCGGAGCGAGGTACTCTACTTCAACCAACATG 3309
DB 181 GACAGCCAGGGCCATCGTCTCAACCGGAGCGAGGTACTCTACTTCAACCAACATG 240

QY 3310 CAGGACCGGGCAGCCAAAGATCGAACCGCGAGCCCTTGGACCGCAGCGAGGTCTC 3369
DB 241 CAGGACCGGGCAGCCAAAGATCGAACCGCGAGCCCTTGGACCGCAGCGAGGTCTC 300

QY 3370 TTCACACCGGCTCATCGGCCCTTGGCCCTGGTGGTGGACACACACTGGGCAAGCTG 3429
DB 301 TTCACACCGGCTCATCGGCCCTTGGCCCTGGTGGTGGACACACACTGGGCAAGCTG 360

QY 3430 TTCTGGTGGACCGGACTGAAGCGCATTTGAGAGCTGTGACCTCTGAGGGGCCAACCGC 3489
DB 361 TTCTGGTGGACCGGACTGAAGCGCATTTGAGAGCTGTGACCTCTGAGGGGCCAACCGC 420

QY 3490 CTGACCTTGGAGGACGCCAAATCGTGGACCTCTTGGGCTTGACCATCTTGGCAAGCAT 3549
DB 421 CTGACCTTGGAGGACGCCAAATCGTGGACCTCTTGGGCTTGACCATCTTGGCAAGCAT 480

QY 3550 CTCTACTGATCGACCGCCAGCAGCATGATCGAGCGTGTGGAGAGACACACCGGGAC 3609
DB 481 CTCTACTGATCGACCGCCAGCAGCATGATCGAGCGTGTGGAGAGACACACCGGGAC 540

QY 3610 AAGCGACTTCGATCCAGGGCGGTGTGCGCCACCTCACTGGCATCCATGCGAGTGAGGAA 3669
DB 541 AAGCGACTTCGATCCAGGGCGGTGTGCGCCACCTCACTGGCATCCATGCGAGTGAGGAA 600

QY 3670 GTCAGCTGGAGGAGTTCAGCCACCATGTGCGCCGTGACAAATGGTGGCTGCTCCAC 3729
DB 601 GTCAGCTGGAGGAGTTCAGCCACCATGTGCGCCGTGACAAATGGTGGCTGCTCCAC 660

QY 3730 ATCTGTATTGCAAGGGTGTGGACACCAACCGTGTCTATGCCAGTCCACCTCGTGTCT 3789
DB 661 ATCTGTATTGCAAGGGTGTGGACACCAACCGTGTCTATGCCAGTCCACCTCGTGTCT 720

QY 3790 CTGCAAGACTCTGTGACCTGTGGAGAGCGCCCACT - GCTCCCGGACCAATTTGCAATG 3848
DB 721 CTGCAAGACTCTGTGACCTGTGGAGAGCGCCCACT - GCTCCCGGACCAATTTGCAATG 780

QY 3849 TGCC - ACAGGGAGATCGACTGTATCCCC - GGGGCTTGGCTGTGACGCTTT - CCC 3903
DB 781 TGCC - ACAGGGAGATCGACTGTATCCCC - GGGGCTTGGCTGTGACGCTTT - CCC 840

QY 3904 GAGTGGATGACCA 3917
DB 841 GAGTGGATGACCA 854
```

RESULT 15
AY407796

```
LOCUS      AY407796      5400 bp      DNA      linear      GSS 12-DEC-2003
DEFINITION Homo sapiens HCM3010 gene, VIRTUAL TRANSCRIPT, partial sequence,
              genomic survey sequence.
ACCESSION   AY407796
VERSION     AY407796.1 GI:39763767
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 5400)
AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
              Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
              Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
              Adams,M.D. and Cargill,M.
              Adams,M.D. and Cargill,M.
              Inferring nonneutral evolution from human-chimp-mouse orthologous
              gene trios
              Science 302 (5652), 1960-1963 (2003)
JOURNAL     14671302
PUBMED      14671302
REFERENCE   2 (bases 1 to 5400)
AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
              Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
              Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
              Adams,M.D. and Cargill,M.
              Direct Submission
              Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
              Rockville, MD 20850, USA
              This sequence was made by sequencing genomic exons and ordering
              them based on alignment.
              Location/Qualifiers
FEATURES             source
                     1..5400
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     <1..>5400
                     /locus_tag="HCM3010"
                     gene
ORIGIN
Query Match      15.4%; Score 781.6; DB 9; Length 5400;
Best Local Similarity 54.1%; Pred. No. 1.2e-159; Indels 69; Gaps 15;
Matches 1957; Conservative 0; Mismatches 1594;

QY 147 GCCGCTCTCTGCTATTGTCACACCGCCGGGAGCTACGGCTGGTGACCGCCGGCGAGTCAA 206
DB 999 GCCTGTGCTGCTGTTCCCAATCGCATCGCAGGTGCTGCC-----ACACCG 1052

QY 207 GCTGGAGTCCACCATCGTGTGTCAGCGGCTGGAGAGTCCGCCCGCAGTGGATTCAGTT 266
DB 1053 CTCTGAGTACACACTGCTGCTTAAACAACCTGGAGAATGCCATTGCCCTTGATTTCCACCA 1112

QY 267 TTCCAAGGAGCGCTGTACTGGACAGACGTGAGCGAGGAGCCATCAAGCAGACCTACTCT 326
DB 1113 CGCGCGAGCTGTTCTTCTGTGTGATGTACATCCCTGGACCGGATCTCCGTGCCAATCT 1172

QY 327 GAACCCAGAGCGGGCGCCCGTGCAGAAAGTGGTCACTCCGGCTGTGTTCTTCCCGAGCG 386
DB 1173 CAAC-----GGCAGCAACGTGGAGGAGTTGTCTACTGGGCTGGAGAGCCCGAGGGG 1226

QY 387 CTTGCTGCTGCACTGGGTGGGCAAGAGCTGTACTGGACGAGCTCAGAGACCAACCGCAT 446
DB 1227 CTTGCTGCTGATTTGGGTGCCATGACAACTCTACTGGACCGACTCAGGCACCTCGAGAT 1286

QY 447 CGAGTGGCCCAACTCAATGGCACAATCCCGAAGGTGCTCTTCTGGCAGGAGCTTTGACCA 506
DB 1287 TGAGTGGCCCAATCTGGACGGGGCCACCGGAGGTGTGCTGTGGCAGAACTTGAGAA 1346

QY 507 GCCTAGGGCCATCGCTTTGGACCCCGCTCA CGGGTACATGTACTGGACAGACTGGGGTGA 566
DB 1347 GCGCCGGGCCATTCCTTTGCATCCCATGGAGGGTACCAATTTACTTGGACAGACTGGGGCAA 1406

QY 567 GACGCGCGGATTCAGCGGGCAGGAGTGGATGGCAGCACCCCGAAGATCATTTGTGACTC 626
DB 1407 CACCCCGGCTATTGAGGCTTCAGCATGGATGGCTCTCGACGCCGCATCATTTGCCGATAC 1466
```

QY 627 GGACATTTACTGGCCCAATGACTGACCATCGACCTGGAGGAGCAGAAGCTCTACTGGGC 686
DB |||||
DB 1467 CCATCTCTTCTGGCCCAATGACTGACCATCGACCTGGAGGAGCAGAAGCTCTACTGGGC 1526
QY 687 TGACGCCAAGCTCAGCTTTCATCCACCGTGCACCACTGGACGGCTGTTTCCGCGAAGAGGT 746
DB |||||
DB 1527 GGATGCTAAGCACCACTGATCATCGAGAGGGCCAACTCTGGATGGGAGTCAACCGCTAAGGCTGT 1586
QY 747 GGTGGAGGGCAGCTGACGCAACCCCTTGCCTGAGCTGACGCTCTCCGGGAGACATCTGTACTG 806
DB |||||
DB 1587 CATTAGCCAGGGCTTCCCGCATCCCTTCCGCTACACAGTGTGTAAGACACGCTGTACTG 1646
QY 807 GACAGACTGGCAGACCCGCTCCATCCATGCTGCAACCAAGCGCACTGGGGGGAAGAGAA 866
DB |||||
DB 1647 GACAGACTGGCAGACCCAGAGACATCAATAGCGCTAAACAAATTTACGGGGAAGACCGAGGA 1706
QY 867 GGAGATCTGTAGTGCCCTCTACTCAACCCATGGAATCCAGGTGCTGAGCCAGGAGCGGCA 926
DB |||||
DB 1707 AATCAVTCGCAACAACTCCCACTTCCCTATGGAATCCACACCTTGCAACCCCGAGCGCA 1766
QY 927 GCCTTTCTTCCACACTCGCTGTGAGGAGGACAAATGGCGGCTGCTCCACCTGTGCTGCT 986
DB |||||
DB 1767 ACCTGCGAGGGAACACCGCTGTGGGGAACAACAGGAGGCTGCAACGCACTGTGTCTG-- 1824
QY 987 GTCCCCAAAGCGAGCTTTCTACATATGCGCTGCCCCCAGCGGTGTGACGTGCAGGACAA 1046
DB |||||
DB 1825 ----CCAGTGGCAGAACTACACTGTGCTGCTGCCCCACTGGCT--CCGCAAGATCAG 1877
QY 1047 CGGAGGAGCTGTAAAGCAGAGCGGAGAGGTGCTGCTGTGGCCCGGCGGACCGACCT 1106
DB |||||
DB 1878 CAGCCACGCTGTGCCAGAGTCTTGACAACTTCTGCTGCTTTTGGCCGGAAGATGGACAT 1937
QY 1107 ACAGAGATCTCGCTGAGNACGCGGAGCTTCAACGACATCTGCTGACAGGTGGAGACAT 1166
DB |||||
DB 1938 CCGTCGAATCAGCTTTGACACAGAGGACCTGTCTGATGATGTCTATCCCACTGGCTGACGT 1997
QY 1167 CCGGCAACGCCATTCGATCGACTAGACCCGCTAGAGGGCTATGTCTACTGGACAGATGA 1256
DB |||||
DB 1998 GCGCAGTGTCTGCGCCCTTGACTGGGACTCCCGGATGACACATGTTACTGGACAGATGT 2057
QY 1227 CGAGGTGCGGCCATTCGCGAGGCGGTACTCTGAGCGGGTCTCGGGCGCAGACGCTGGTCAA 1286
DB |||||
DB 2058 CAGCACTGATACCATCAGCAGGGCCAGTGGGATGGAAACAGGAGGTGGTAGTGA 2117
QY 1287 CACCGAGATCAACGACCCGATGCGATCGCGGTGCACTGGGTGGCCCGGAACCTCTACTG 1346
DB |||||
DB 2118 TACCAGTTTGGAGAGCCAGCTGGCCCTGGCCATTGATTTGGGTCAACCAAACTGTACTG 2177
QY 1347 GACCGACACGGGCGACGGACCGCATCGAGGTGACGGCTCAACGGCACCTCCCGCAAGAT 1406
DB |||||
DB 2178 GACAGATGACAGTACAGACCGGATTAAGTAGCCNAACACAGATGGCAGCATGAGAACGT 2237
QY 1407 CCTGGTGTGCGAGACCTTGAACGAGCCCGAGCCATGCACTGCAACCCCGTGTATGGGCT 1466
DB |||||
DB 2238 ACTCATCTGGGAGAACCTTGTATCGTCTCGGACATCGTGTGGAAACCATGGGCGGTA 2297
QY 1467 CATGTACTGGACAGACTGGGAGAGAACCTTAATCGAGTGTGCCAACTTGGATGGGCA 1526
DB |||||
DB 2298 CATGTATTGGACTGACTGGGTGGAGCCCAAGATTGAAACGAGCTGGCATGGATGCCCTC 2357
QY 1527 GGAGCGGCTGTGTGTTCAATGCTCCTCGGTGGGCCCAACGGCTTGGCCCTGGACCT 1586
DB |||||
DB 2358 AGGCGCCAACTGATATTCTTCTAATCTGACCTGGGCTTAATGGGTAGCTATTGATTA 2417
QY 1587 GCAGAGGGGAGCTCTACTGGGAGAGCCCAAGACAGACAGATTCGAGGTGATCAATGT 1646
DB |||||
DB 2418 TGGGTCCAGCGTCTACTTGGCTGACGCGCGCATGAAGACAATTTGAATTTGTGGACT 2477
QY 1647 TGATGGGAGGAAGAGGGGAGCCCTCTGAGGAGACAGCTCCGCGACATTTTCGGGTTTCA 1706
DB |||||
DB 2478 GGATGGCAGTAAAGAGGAGGTGCTGATTTGGAAGCAGCTCCCGACCACTTTCGGCTGAC 2537

QY 1707 GCTGCTGGGGAGCTTTCATCTACTGGAAGTCTGCTGCGAGCGCGCAGCATCGAGCGGTGCA 1766
DB |||||
DB 2538 CCTCTATGGAGAGCGCATCTATTGGAAGTCTGCTGCGAGCAACCAAGAGCATACAGAGCGCTGA 2597
QY 1767 CAAGGTCA---AGGCCAGCGGGAGCGTCACTATTGACAGCTGCCCGGACCTGTATGGGGCT 1823
DB |||||
DB 2598 CCGCTGACAGGGCTGGACCGGAGACTCTGTCAGGAGAACTTGGAAACCTTAATGGACAT 2657
QY 1824 CAAGAGCTGTGAATGTGGCCAGGTCTGCGGAACCAACCCGCTGCGGACAGGACAGGCGGG 1883
DB |||||
DB 2658 CCATGTCTTCAACGCGCGCGGCCCCAGTGTCTACACATGTGTATGGAGAAATGGCGG 2717
QY 1884 GTGAGCCACCTGT-----GCTTTTTCACACCCCAACCGGCTGGGTGGCTGCCCAT 1937
DB |||||
DB 2718 CTGTAGCCACCTGTGTCTTAGTCCCAAAATCCAAGCGGATTGAGTGTACTGTGCCAC 2777
QY 1938 CGGCTGGAGCTGTGTAGTGTGACATGAAGACCTGC---ATCGTGCCTGAGGGCTTCTTGGT 1994
DB |||||
DB 2778 AGGCATCAACCTGTGTCTGATGGCAAGACCTGTCTACAGGCGATGAACAGTTCCTCAT 2837
QY 1995 CTTCCACAGCAGAGCGCGCATCCACAGATCTCCTCGAGACCAATAA-----CAACGA 2048
DB |||||
DB 2838 CTTCCGAGGAGGATAGACATTCGATGTCTCCTCGACATCCCTTATTTTGTGTGATGT 2897
QY 2049 CGTGGCCATCCCGCTCACGGGCGTCAAGGAGGCTCTCAGCCCTGGACTTTGATGTCTCAA 2108
DB |||||
DB 2898 GGTGTACCAATCAACATTACCATGAAGACACCATTTGCCGTTGGAGTAGACCCCGAGGA 2957
QY 2109 CAACCAATCTACTGGACAGACGTGACCTGAAGACCATCAGCCGCGCTTTCATGAACGG 2168
DB |||||
DB 2958 AGGAAAGGTGTACTGTCTGACAGCACATGACAGGATCAGTGTGTCATCTGCGATG 3017
QY 2169 GAGCTCGCTGGAGCACGCTGTGGAGTTTGGCTTGTACTACCCCGAGGCGCATGGCGTTGA 2228
DB |||||
DB 3018 CTCACAGATGAGGACATCATCACAGGGCTACAGACACAGATGGGCTCGCGGTTGA 3077
QY 2229 CTGATGGGGAAGAACCTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTGGCGG 2288
DB |||||
DB 3078 TGCCATTTGGCGGAAAGTATCTGGAACAGACACCGGAACAAACCGGATTGAAGTGGGCA 3137
QY 2289 GCTGGAAGGGAGTTCGCGCAAGTCTCTGTGTGAGGAGCTTGGACAAACCCGAGGTCTGT 2348
DB |||||
DB 3138 CCTGGACGGGTCCATGCGGAAAGTGTGGTGTGCGAGAACCTTGACAGTCCCGCGGCCAT 3197
QY 2349 GGCCCTCGATCCCAACAGGCTTACTCTGACCGAGTGGGGCGGCGCAACCGAGAT 2408
DB |||||
DB 3198 CGTACTGTACATGAGATGGGGTTTATGTACTGACAGACTGGGGGGAGAAATGCCAAGTT 3257
QY 2409 CGTGGGGCTTCAATGGACGGGACCAACTGATGACGCTGGT---GGACAAAGTGGGCGG 2465
DB |||||
DB 3258 AGAGCGTCCGGAATGATGGCTCAGACCGCGCGGTCTCATCAACAAACCTAGGATG 3317
QY 2466 GGCCAAAGCCTCAACATTTGATGCTGACGCTGACGAGGCTTCTACTGGACCGACCTGGGAC 2525
DB |||||
DB 3318 GCCCAATGGAATGCTGTGGCAAGGCGAGCTCCCAACTGTCTATGGGCGGATGCGCCAC 3377
QY 2526 CAACATGATCGAGTCGTCCAACTGCTGGGTGAGGAGGGGTGCTGATTTGCCAGCATCT 2585
DB |||||
DB 3378 CGAGCAATTTGAGGCTGTGACCTGAAATGGTGGTCCAAATCGGCATATCATTTGGTGTCA 3437
QY 2586 CCCGCAACCGGTTCCGCTGACGCGAGTACAGCGATTTATATCTACTGGACAGACTGGAAATCT 2645
DB |||||
DB 3438 GCAGCAACCATATGCGCTTCAACCTGCTGCACTTCTATCTACTGGACTGATGCGACAG 3497
QY 2646 GCAAGCAATTTAGCGGGCGGCAAGACTAGCGGCGGAAACCGACCTTCATTCAGGGGCA 2705
DB |||||
DB 3498 TCGGAGCATCCACCGTGTCAAGAGGTACTGGCAGCAATGTCTCTCTGTGAGTCCAA 3557
QY 2706 CTTGGAATTTGTGATGACATCTGTGTTCCTCTCTCCCGCAGGATGCGCTCAATGA 2765
DB |||||
DB 3558 CTTCCAGGCTCTATGACATGCGAGGTGTGGACCGGGGACAGCACTAGGTGTTTAAACA 3617
QY 2766 CTGTATGCAACAAACGGGCGAGTGTGGCGAGCTGTGCTTGGCATCTCCCGCGGCGCACCG 2825

Db	3618	GTGCGGCTCAGAGAAATGGCGGCTGCTCCCACTCTGCTTGCCCTGGCGCCTTCTGGCTTCTC	3677
Qy	2826	CTGCGGCTGGCCCTCACTACACCTCTGGACCCAGCAGCGCAACTGACGCGCGCCAC	2885
Db	3678	CTGTGCTCTCCCACTGGCATCCAGCTGAAGGAGATGGGAAGACCTGTGTATCCCTCTCC	3737
Qy	2886	C----ACCTCTTGTGTTTCAGCAGAAATCTGCCATCAGTCGATGATCCCGACACCA	2942
Db	3738	TGAGACCTACTGCTCTTCTTCAGCGCGTGCTCCATCCCGCGTATCTCATGTGACACCA	3797
Qy	2943	GCACGCGCGGATCTCATCTCTGCCCTGCAATGGACTGAGGAAGTCAAAAGCCATCGACTA	3002
Db	3798	TGACCAACCGATGTGCATGTCCCTGTTCTTGAGCTCAACATGTCTATCTCCCTGACTA	3857
Qy	3003	TGACCCACTGGACAAAGTTTCACTCTGGGTGGATGGCGCGCAGAAATCAAGCGAGCCAA	3062
Db	3858	TGACAGCGTGGATGGAAGGCTATTACACAGATGTGTTCTCGATGTTA--TCAGGCGA	3915
Qy	3063	GGACGCGGACCCAGCCCTTTGTTTGAACCTCTCTGAGCGAAGGCCAAACCCAGACAG	3122
Db	3916	GCAGCTGAAACCGCAGCAACATGGAGACAGTGAATCGGGCGAGGGCTGAAGACA	3970
Qy	3123	GCAGCCCAAGCACTCAGACATCGACATCTACAGCGGACACTGTCTTCTGGACGTGCGAGG	3182
Db	3971	-----CTGACGGCTGCGACGTGGACTGGGTGGCGAGGAACCTGTACTGGACACACAG	4025
Qy	3183	CACCAATACATCAAGTCCACAGGCTGACGGGGGAAGCCATGGGGGTGTGCTGCTGG	3242
Db	4026	TCGAAATACCAATGAGCGCTCCAGGCTGGATGTTCTGCGCAAGTACTGATCAACAA	4085
Qy	3243	GGACGGCAACAAAGCCAGGCGCAATCGTCGTCACCGCGAGCGAGGTACCTGTACTTTCAC	3302
Db	4086	TAGCTTGATGAGCCCGCGGCCATTGCTGTTTCCCAGGAAGGGGTACTCTTCTGGAC	4145
Qy	3303	CAACATCGAGAACGGCGAGCAAGATCGAAACGCGACCGCTTGGACGGCAACGAGCGCGA	3362
Db	4146	AGAC---TGGGGCCACATTGCAAGATCGAAACGGGCAAACTTGGATGGTTCTGAGCGAA	4202
Qy	3363	GGTCCTCTTACCAACCGGCTCATCCGCCCTGTGGCCCTGGTGTGGACAAACACTGGG	3422
Db	4203	GGTCTCATCAACAGACCTTGGTTGGCCCAATGGCCTTTACCTTGAATGATATCCCG	4262
Qy	3423	CAAGCTGTTCTGGGTGGAACGGCACTTGAAGCGCATTTGAGAGCTGTGAACCTGTGAGGGC	3482
Db	4263	CAGGATCTACTGGGTGGATGCGCATCTGGACCGGATCGAGAGTGTGACCTCAATGG--	4320
Qy	3483	CAACCGCTGACCTTGAAGGACGCAACATGTGACGCTCTGGGCTGACCATCTTGG	3542
Db	4321	-AAACTCGGAGGTCTTGTGCGGCCATGTGTCCACCCCTTTGGCCCTCACAGCAAGA	4379
Qy	3543	CAAGCATCTTACTTGAATCGACCGCCAGCAGCAGATGATCGAGCGTGTGGAGAAGACCAC	3602
Db	4380	CAGTGGATCTACTTGGACAGACTGGCAGACCAAGTCAATCCAGGCTGTTGACAAATCTC	4439
Qy	3603	CGGGGAACAAGCGGACTCGCATCCAGGCGCGTGTGCCCAACTCATCTGGCATTCATGCAGT	3662
Db	4440	AGGCCGGAAACAAAGGAGACAGTGTCTGGCAAAATGTGGAAAGGACTCATGATATC-----AT	4493
Qy	3663	GGAGGAAGTCAAGCTTGAGGAGTTCTCAGGCCCAACCATGTGCCGTGACAATGTGGCTG	3722
Db	4494	CGTGGTTCCCTCAGCGGCAGCAGGGAACCAATGCCCTGTGGTGNAACAAATGGTGGCTG	4553
Qy	3723	CTCCCACTCTGTAATGCCA	3742
Db	4554	CACCCACTCTGTTTTGCCA	4573

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2005, 12:03:25 ; Search time 2493 Seconds
(without alignments)
12022.323 Million cell updates/sec

Title: US-09-931-375A-1
Perfect score: 5063
Sequence: 1 gccatggagcccgagtgagc.....aggctgggagaactttgta 5063

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5063	100.0	5063	6 AAD34322	Aad34322 Human bon
2	5061.4	100.0	5063	6 AAD34405	Aad34405 Human BSM
3	5061.4	100.0	5063	6 AAD34401	Aad34401 Human BSM
4	5061.4	100.0	5063	6 AAD34404	Aad34404 Human BSM
5	5058.2	99.9	5120	5 ABA82617	Abas2617 Human Zma
6	5058.2	99.9	5120	8 ACC45358	Acc45358 Human wil
7	5058.2	99.9	5120	10 ADB99038	Adb99038 Human LRP
8	5058.2	99.9	5120	10 ADE82639	Ade82639 Human Zma
9	5058.2	99.9	5120	13 ADRI6919	Adri6919 Human hig
10	5058.2	99.9	5278	8 ACC46079	Acc46079 Human ZMA
11	5058.2	99.9	5278	10 ADB98797	Adb98797 Human ZMA
12	5056.8	99.9	5099	13 ABD32937	Abd32937 Human can
13	5056.6	99.9	5120	8 ABK22777	Abk22777 Human hig
14	5056.6	99.9	5120	8 ACC45359	Acc45359 Human LRP
15	5056.6	99.9	5120	10 ADB99039	Adb99039 LRP5 vari
16	5056.6	99.9	5120	10 ADE82640	Ade82640 Human HBM
17	5056.6	99.9	5120	13 ADRI6920	Adri6920 Human hig
18	5056.6	99.9	5278	8 ACC46078	Acc46078 Mouse HBM
19	5056.6	99.9	5278	10 ADB98796	Adb98796 Mouse HBM
20	5055.2	99.8	5099	13 ADQ86437	Adq86437 Human tum

21	5055.2	99.8	5099	13 ADQ83350	Adq83350 Human tum
22	5055	99.8	5120	5 ABA82618	Abas2618 Human HBM
23	5052	99.8	5100	10 ADK66938	Adk66938 Gene #28
24	5052	99.8	5100	13 ADR73481	Adr73481 Human low
25	5050	99.7	5066	6 AAD34403	Aad34403 Human BSM
26	5036.6	99.5	5112	13 ADRA47571	Adra47571 Human hig
27	5032.2	99.4	5115	13 ADRA47570	Adra47570 Human hig
28	5028	99.3	5098	2 AAV70396	Aav70396 LRP5 isof
29	5010.8	99.0	5135	11 ACN44787	Acn44787 Human mRN
30	5010.8	99.0	5135	13 ABD32938	Abd32938 Human can
31	4999	98.7	5059	12 ADQ20523	Adq20523 Human sof
32	4952.6	97.8	5599	12 ADQ24509	Adq24509 Human sof
33	4902	96.8	5022	2 AAV85549	Aav85549 LRP5 isof
34	4894.4	96.7	5263	2 AAV70400	Aav70400 LRP5 isof
35	4890.6	96.6	5125	2 AAV86019	Aav86019 LRP5 isof
36	4890	96.6	5165	2 AAV85551	Aav85551 LRP5 isof
37	4889.6	96.6	5162	2 AAV85550	Aav85550 LRP5 isof
38	4840	95.6	4848	12 ADI27199	Adi27199 Human LRP
39	4813	95.1	4843	2 AAV70395	Aav70395 LRP5 CDNA
40	4723.6	93.3	4915	2 AAV70398	Aav70398 LRP5 isof
41	4584.2	90.5	4736	6 ABK22776	Abk22776 Human Zma
42	4322.6	85.4	4351	2 AAV70397	Aav70397 LRP5 isof
43	3931	77.6	5119	6 ABA94317	Abas4317 LRP5 poly
44	3907	77.2	5117	2 AAV86018	Aav86018 Mouse Lrp
45	3900.2	77.0	5018	13 ABD32935	Abd32935 Mouse can

ALIGNMENTS

RESULT 1
AAD34322
ID AAD34322 standard; cDNA; 5063 BP.
XX
AC AAD34322;
XX
DT 16-JUL-2002 (first entry)
XX
DE Human bone strength and mineralisation regulatory protein (BSMR) CDNA.
XX
KW Human; bone strength and mineralisation regulatory protein; BSMR;
KW bone strength; mineralisation; ophthalmological; antidiabetic;
KW bone density regulating transmembrane receptor; prosthetic device;
KW surgical implant; diabetic retinopathy; hypertensive retinopathy;
KW therapy; osteoporosis; prematurity; ocular vessel; eye disorder;
KW osteopathic; Gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 52..4899
FT /*tag= a
FT /product= "Human BSMR protein"
XX
XX WO200216553-A2.
XX
XX 28-FEB-2002.
XX
XX 17-AUG-2001; 2001WO-US041788.
XX
XX 18-AUG-2000; 2000US-0226119P.
XX
XX 22-SEP-2000; 2000US-0234337P.
XX
XX 13-JUL-2001; 2001US-0304851P.
XX
XX (AVET) AVENTIS PHARMA SA.
XX
XX (HARD) HARVARD COLLEGE.
XX
XX (UYCA-) UNIV CASE WESTERN RESERVE.
XX
XX Warman ML, Gong Y, Olsen BR, Rawadi G, Roman-Roman S;
XX
XX WPI; 2002-329694/36.
XX
XX P-PSDB; AAE21730.

QY 361 ATCTCCGGCTGTCTCTCCCGACGGGCTCGCTGCGATGCGGTGGGCAAGAGCTGTAC 420
DB 361 ATCTCCGGCTGTCTCTCCCGACGGGCTCGCTGCGATGCGGTGGGCAAGAGCTGTAC 420
QY 421 TGGACGAGCTCAGAGACAAACCGCATCGAGGTGGCCAACTCAATGGCACATCCCGAAG 480
DB 421 TGGACGAGCTCAGAGACAAACCGCATCGAGGTGGCCAACTCAATGGCACATCCCGAAG 480
QY 481 GTGCTCTTTCTGGCAGGACCTTGAACAGCTTAGGCCATTCGCTTGGACCCCGCTCAACGG 540
DB 481 GTGCTCTTTCTGGCAGGACCTTGAACAGCTTAGGCCATTCGCTTGGACCCCGCTCAACGG 540
QY 541 TACATGTACTGGACAGACTGGGTGAGACGCGCCGGATTGAGCGGGCAGGATGGATGGC 600
DB 541 TACATGTACTGGACAGACTGGGTGAGACGCGCCGGATTGAGCGGGCAGGATGGATGGC 600
QY 601 AGCACCCGGAAGATCATTTGTGACTCGGACATTTTACTGGCCCAATGGACTGACCATCGAC 660
DB 601 AGCACCCGGAAGATCATTTGTGACTCGGACATTTTACTGGCCCAATGGACTGACCATCGAC 660
QY 661 CTGGAGGAGCAAGACTCTACTTGGGCTGACGCAAGCTCAGCTTCATCCACCGTGCCAAAC 720
DB 661 CTGGAGGAGCAAGACTCTACTTGGGCTGACGCAAGCTCAGCTTCATCCACCGTGCCAAAC 720
QY 721 CTGGACGGCTGTTCGCGCAGAGGTGTGGAGGCGAGCTGACGACACCCCTTGGCCCTG 780
DB 721 CTGGACGGCTGTTCGCGCAGAGGTGTGGAGGCGAGCTGACGACACCCCTTGGCCCTG 780
QY 781 ACGCTCTCCGGGACACTCTGTACTGAGACAGACTGGCAGACCCGCTCCATCCATCGCTGC 840
DB 781 ACGCTCTCCGGGACACTCTGTACTGAGACAGACTGGCAGACCCGCTCCATCCATCGCTGC 840
QY 841 AACAAAGCCTCTGGGGGGAAGAGAGATCTCTGAGTGCCCTTACTCAACCCATGGAC 900
DB 841 AACAAAGCCTCTGGGGGGAAGAGAGATCTCTGAGTGCCCTTACTCAACCCATGGAC 900
QY 901 ATCCAGGTGCTGAGCAGGAGGCGGAGCTTTCTTCCACACTCGCTGTGAGGAGCAAT 960
DB 901 ATCCAGGTGCTGAGCAGGAGGCGGAGCTTTCTTCCACACTCGCTGTGAGGAGCAAT 960
QY 961 GCGGCTGTCTCCACTGTGCTGTCTCTCCCAAGCGAGCCCTTCTACACATGCGCCCTGC 1020
DB 961 GCGGCTGTCTCCACTGTGCTGTCTCTCCCAAGCGAGCCCTTCTACACATGCGCCCTGC 1020
QY 1021 CCCACGGGTGTGACGTGACGAGCAAAACGCGAGCAAGTGTAAAGGACGAGGAGGAGGTG 1080
DB 1021 CCCACGGGTGTGACGTGACGAGCAAAACGCGAGCAAGTGTAAAGGACGAGGAGGAGGTG 1080
QY 1081 CTGCTGTGCGCGGCGGACGACCTACGGAGATCTCGCTGGACACGCGGACTTCACC 1140
DB 1081 CTGCTGTGCGCGGCGGACGACCTACGGAGATCTCGCTGGACACGCGGACTTCACC 1140
QY 1141 GACATCGTGTGCTGAGGTGACGACATCCGGCAGCGCATTTGCCATCGACTACGACCCGCTA 1200
DB 1141 GACATCGTGTGCTGAGGTGACGACATCCGGCAGCGCATTTGCCATCGACTACGACCCGCTA 1200
QY 1201 GAGGCTATGTCTATGGACAGATGACGAGGTGGGGCCATCCGAGGCGGTACCTGGAC 1260
DB 1201 GAGGCTATGTCTATGGACAGATGACGAGGTGGGGCCATCCGAGGCGGTACCTGGAC 1260
QY 1261 GGGTCTGGGCGCAGACGCTGTCTAACACCGAGATCAACGACCCCGATGGCATCCGGTTC 1320
DB 1261 GGGTCTGGGCGCAGACGCTGTCTAACACCGAGATCAACGACCCCGATGGCATCCGGTTC 1320
QY 1321 GACTGGGTGGCCCGAAACCTCTACTGGACCGACGAGGACGAGCCGATCGAGGTGAGC 1380
DB 1321 GACTGGGTGGCCCGAAACCTCTACTGGACCGACGAGGACGAGCCGATCGAGGTGAGC 1380
QY 1381 CCGCTCAACGGCACTCTCCGCAAGATCTTGTGTGAGGACCTGGAAGCCCGGAGCC 1440
DB 1381 CCGCTCAACGGCACTCTCCGCAAGATCTTGTGTGAGGACCTGGAAGCCCGGAGCC 1440

QY 1441 ATCGCACTGCACCCCGTATGGGCTCATGTACTGGACAGACTGGGAGAGAAACCTTAAA 1500
DB 1441 ATCGCACTGCACCCCGTATGGGCTCATGTACTGGACAGACTGGGAGAGAAACCTTAAA 1500
QY 1501 ATCGAGTGTGCAAACTTTGGATGGGACGAGCGGTGTGTGGTCAATGCCCTCCCTCGG 1560
DB 1501 ATCGAGTGTGCAAACTTTGGATGGGACGAGCGGTGTGTGGTCAATGCCCTCCCTCGG 1560
QY 1561 TGGCCCAACCGGCTGGCCCTGGACCTCGAGAGGGGAGCTCTACTGGGAGAGCCCAAG 1620
DB 1561 TGGCCCAACCGGCTGGCCCTGGACCTCGAGAGGGGAGCTCTACTGGGAGAGCCCAAG 1620
QY 1621 ACACAGAATTCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCTCTCTGGAGAC 1680
DB 1621 ACACAGAATTCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCTCTCTGGAGAC 1680
QY 1681 AAGCTCCCGCAATTTTTCGGGTTTCACTGCTGGGGGACTTCTACTTGAAGTACTGG 1740
DB 1681 AAGCTCCCGCAATTTTTCGGGTTTCACTGCTGGGGGACTTCTACTTGAAGTACTGG 1740
QY 1741 CAGCGCCGACGATCGAGCGGGTGCAAGGTCAAGGCCAGCGCGGAGGTCTCATTTGAC 1800
DB 1741 CAGCGCCGACGATCGAGCGGGTGCAAGGTCAAGGCCAGCGCGGAGGTCTCATTTGAC 1800
QY 1801 CAGCTGCCCGACCTGATGGGCTCAAAAGCTGTGAATGTGGCCAAAGGTCTCGGAAACCAAC 1860
DB 1801 CAGCTGCCCGACCTGATGGGCTCAAAAGCTGTGAATGTGGCCAAAGGTCTCGGAAACCAAC 1860
QY 1861 CCGTGTGCGGACAGGAAACGGGGGTGAGCCACTGTGCTTTTCAACCCCAACCAAC 1920
DB 1861 CCGTGTGCGGACAGGAAACGGGGGTGAGCCACTGTGCTTTTCAACCCCAACCAAC 1920
QY 1921 CCGTGTGCGTCCCATCGGCTGGAGCTGTGATGATGATGATGATGATGATGATGATG 1980
DB 1921 CCGTGTGCGTCCCATCGGCTGGAGCTGTGATGATGATGATGATGATGATGATGATG 1980
QY 1981 GAGGCTTTCTTGGTCTTCCACGAGAGCGCCCATCCACAGGATCTCTCCCTCGAGACCAAT 2040
DB 1981 GAGGCTTTCTTGGTCTTCCACGAGAGCGCCCATCCACAGGATCTCTCCCTCGAGACCAAT 2040
QY 2041 AACAAAGCAGTGGGCTATCCCGCTCAAGGCGCTCAAGGAGGCTCAGCCCTGGATTTGAT 2100
DB 2041 AACAAAGCAGTGGGCTATCCCGCTCAAGGCGCTCAAGGAGGCTCAGCCCTGGATTTGAT 2100
QY 2101 GTGTCCAAACCAACATCTACTGGAAGAGCTGAGCTGAGGAGGCTCAGCCCTGGATTTGAT 2160
DB 2101 GTGTCCAAACCAACATCTACTGGAAGAGCTGAGCTGAGGAGGCTCAGCCCTGGATTTGAT 2160
QY 2161 ATGAAACGGGAGCTCGGTGGAGCAGCTGTGGAGTTTGGGCTTGAATACCCCGAGGCGATG 2220
DB 2161 ATGAAACGGGAGCTCGGTGGAGCAGCTGTGGAGTTTGGGCTTGAATACCCCGAGGCGATG 2220
QY 2221 GCGGTTGATCGGATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGATCGAA 2280
DB 2221 GCGGTTGATCGGATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGATCGAA 2280
QY 2281 GTGGCGGCTGGAGCGGCGAGTCCGCAAGTCTCTGTGTGGAGGACTTGGACAAACCG 2340
DB 2281 GTGGCGGCTGGAGCGGCGAGTCCGCAAGTCTCTGTGTGGAGGACTTGGACAAACCG 2340
QY 2341 AGGTGCTGGGCTGGATTCACCAAGGCTTACATCTACTGGACCGAGTGGGCGGCGAAG 2400
DB 2341 AGGTGCTGGGCTGGATTCACCAAGGCTTACATCTACTGGACCGAGTGGGCGGCGAAG 2400
QY 2401 CCGAGGATCTGGGCGCTTTCATGGAAGGAGCAACTCTGATGAGCGCTGGTGGAGAGGTG 2460
DB 2401 CCGAGGATCTGGGCGCTTTCATGGAAGGAGCAACTCTGATGAGCGCTGGTGGAGAGGTG 2460
QY 2461 GCGCGGGCAACGAGCTCAACATTTGATCTGCTGACGAGCGCTTACTTGGACCGACCTG 2520
DB 2461 GCGCGGGCAACGAGCTCAACATTTGATCTGCTGACGAGCGCTTACTTGGACCGACCTG 2520
QY 2521 GACACCAACATGATCGAGTCTGTCACCAATGTGGGTGAGGAGCGGCTGTGATTTGCCGAC 2580

Db	2521	GACACCAACATGATCGAGTCTGTCAAATGCTGGGTGAGAGCGGGTCGTGATTGCCGAC	2581
Qy	2581	GATCTCCGGCAGCCCGTTTCGGTCTTGACGCGAGTACAGCGATTATATCTACTTGGACAGACTGG	2640
Db	2581	GATCTCCGGCAGCCCGTTTCGGTCTTGACGCGAGTACAGCGATTATATCTACTTGGACAGACTGG	2640
Qy	2641	AATCTGCACAGCATGTAGAGCGGGCCGACAAGACTAGCGCGCGGAACCGCACTCATCCAG	2700
Db	2641	AATCTGCACAGCATGTAGAGCGGGCCGACAAGACTAGCGCGCGGAACCGCACTCATCCAG	2700
Qy	2701	GGCCACCTGGACTTCGTGATGGAGCATCCTGGTGTTCACCTCTCCCGCAGAGATGGCTC	2760
Db	2701	GGCCACCTGGACTTCGTGATGGAGCATCCTGGTGTTCACCTCTCCCGCAGAGATGGCTC	2760
Qy	2761	AATGACTGTATGCACAAACCGGCGAGTGTGGGCGAGCTGTGCCCTTGGCATCTCCCGGGCGC	2820
Db	2761	AATGACTGTATGCACAAACCGGCGAGTGTGGGCGAGCTGTGCCCTTGGCATCTCCCGGGCGC	2820
Qy	2821	CACGGCTGGGGCTGCGCTCACACTACACCTTGGACCCCGCAGCGCCAACTGCAGCCG	2880
Db	2821	CACGGCTGGGGCTGCGCTCACACTACACCTTGGACCCCGCAGCGCCAACTGCAGCCG	2880
Qy	2881	CCACACACCTTCCTGCTTTCAGCCAGAAATCTGCCATCATGTCCGATGATCCCGGACGAC	2940
Db	2881	CCACACACCTTCCTGCTTTCAGCCAGAAATCTGCCATCATGTCCGATGATCCCGGACGAC	2940
Qy	2941	CAGCACAGCCGGATCTCATCTCTGCCCTGCATCGACTGAGGAACTGCAAGCCATCGAC	3000
Db	2941	CAGCACAGCCGGATCTCATCTCTGCCCTGCATCGACTGAGGAACTGCAAGCCATCGAC	3000
Qy	3001	TATGACCCACTGGACAAGTTCACTACTCGGTGGATGGGCGCCAGAACCAACCCAGAC	3060
Db	3001	TATGACCCACTGGACAAGTTCACTACTCGGTGGATGGGCGCCAGAACCAACCCAGAC	3060
Qy	3061	AAGGACGACGGGACCCAGGCCCTTTGTTTGAACCTCTTGAGCCAGGCGCCAGAACCCAGAC	3120
Db	3061	AAGGACGACGGGACCCAGGCCCTTTGTTTGAACCTCTTGAGCCAGGCGCCAGAACCCAGAC	3120
Qy	3121	AGGCGAGCCCAAGACCTCAGCATCGACTCTACAGCCGAGACACTGTTCTCGAGCTGCGAG	3180
Db	3121	AGGCGAGCCCAAGACCTCAGCATCGACTCTACAGCCGAGACACTGTTCTCGAGCTGCGAG	3180
Qy	3181	GCCACCAATACCATCAACGTCCAGGCTCAGCGGGGAAGCCATGGGGGTGGTGTGCGT	3240
Db	3181	GCCACCAATACCATCAACGTCCAGGCTCAGCGGGGAAGCCATGGGGGTGGTGTGCGT	3240
Qy	3241	GGGGAACCGGACAAAGCCACAGGCCATCGTCTGTAACCGCGAGCGAGGTTACCTGTGTTTC	3300
Db	3241	GGGGAACCGGACAAAGCCACAGGCCATCGTCTGTAACCGCGAGCGAGGTTACCTGTGTTTC	3300
Qy	3301	ACCAACATGACAGACCGGGCAGCCAGATCGAACGGCAGCCCTGAGCGCCACCGAGCGC	3360
Db	3301	ACCAACATGACAGACCGGGCAGCCAGATCGAACGGCAGCCCTGAGCGCCACCGAGCGC	3360
Qy	3361	GAGGTCTCTTTCACACCGGCCCTCATCCGCCCTGTGGCCCTGTGTGGTGGACAAACAATG	3420
Db	3361	GAGGTCTCTTTCACACCGGCCCTCATCCGCCCTGTGGCCCTGTGTGGTGGACAAACAATG	3420
Qy	3421	GGCAAGCTGTTCTGGGTGGACCGGGACCTGAAGCGCATTGAGAGCTGTGACTGTGAGGG	3480
Db	3421	GGCAAGCTGTTCTGGGTGGACCGGGACCTGAAGCGCATTGAGAGCTGTGACTGTGAGGG	3480
Qy	3481	GCCAAACCGCTGACCTTGGAGGACGCAACATCGTGCAGCCTCTGGCCCTGACCATCCCT	3540
Db	3481	GCCAAACCGCTGACCTTGGAGGACGCAACATCGTGCAGCCTCTGGCCCTGACCATCCCT	3540
Qy	3541	GGCAAGCATCTCTACTTGGATCGACCGGCACGACGATGATCGAGCGTGTGGAGAGAGACC	3600
Db	3541	GGCAAGCATCTCTACTTGGATCGACCGGCACGACGATGATCGAGCGTGTGGAGAGAGACC	3600
Qy	3601	ACCGGGACAAAGCGGACTCGCATCCAGGGCCGTGTGGCCCACTCATGTGCACTTCATGCA	3660

[illegible]

Qy	721	CTGACGGCTCGTTTCGGCAGAGAGGTGGTGGAGGGCAGCCTGACGCAACCCCTTCGCCCTG	780
Db	721	CTGACCGGCTCGTTTCGGCAGAGAGGTGGTGGAGGGCAGCCTGACGCAACCCCTTCGCCCTG	780
Qy	781	ACGCTCTCCGGGACACTCTGTACTTGACAGACTGGCAGACCCGCTCCATCCATGCGCTGC	840
Db	781	ACGCTCTCCGGGGACACTCTGTACTTGACAGACTTGGCAGACCCGCTCCATCCATGCGCTGC	840
Qy	841	AACAAGCGCACTGGGGGGAAGAGGAGATCTTGAGTGGCCCTCTACTCACCCATGGAC	900
Db	841	AACAAGCGCACTGGGGGGAAGAGGAGATCTTGAGTGGCCCTCTACTCACCCATGGAC	900
Qy	901	ATCAGAGTGTGACGACAGGAGCGGACGCTTTCTTCCACACTCGCTGTGAGGAGCAAT	960
Db	901	ATCCAGGTGTGACGACAGGAGCGGACGCTTTCTTCCACACTCGCTGTGAGGAGCAAT	960
Qy	961	GGGGCTGCTCCCACTGTGCTGCTGTGCCCAAGCGAGCGTTTCTATACATCGGCTGC	1020
Db	961	GGGGCTGCTCCCACTGTGCTGCTGTGCCCAAGCGAGCGTTTCTATACATCGGCTGC	1020
Qy	1021	CCCAAGGCTGTGACGTGCGAGCAACCGGAGGACGTGTAAAGCAGAGACCGGAGGCTG	1080
Db	1021	CCCAAGGCTGTGACGTGCGAGCAACCGGAGGACGTGTAAAGCAGAGACCGGAGGCTG	1080
Qy	1081	CTGCTGTGCGCCGGCGGACGGACCTTACGAGGATCTCGCTGGACACGCCGACTTCACC	1140
Db	1081	CTGCTGTGCGCCGGCGGACGGACCTTACGAGGATCTCGCTGGACACGCCGACTTCACC	1140
Qy	1141	GACATCGTGTGAGGTGGAGGACATCCGGCACGCCATTCGCATCGACTACGACCCGCTA	1200
Db	1141	GACATCGTGTGAGGTGGAGGACATCCGGCACGCCATTCGCATCGACTACGACCCGCTA	1200
Qy	1201	GAGGCTATGTCTACTGCAACAGATGACGAGTGGCGGCCATTCGACGGGCGTACTTGGAC	1260
Db	1201	GAGGCTATGTCTACTGCAACAGATGACGAGTGGCGGCCATTCGACGGGCGTACTTGGAC	1260
Qy	1261	GGGTCTGGGGCGCAGACGCTGGTCAACAACGAGATCAACACGCCGATGCGCGGTC	1320
Db	1261	GGGTCTGGGGCGCAGACGCTGGTCAACAACGAGATCAACACGCCGATGCGCGGTC	1320
Qy	1321	GACTGGGTGGCCGAACCTCTACTGACCGCACCGGGCAGCACCGCATCGAGTGAAGC	1380
Db	1321	GACTGGGTGGCCGAACCTCTACTGACCGCACCGGGCAGCACCGCATCGAGTGAAGC	1380
Qy	1381	CGCTCAACGGCACCTCCCGCAAGATCCTGGTGTGGAGGACCTTGGACGAGCCCCGAGCC	1440
Db	1381	CGCTCAACGGCACCTCCCGCAAGATCCTGGTGTGGAGGACCTTGGACGAGCCCCGAGCC	1440
Qy	1441	ATCGACTGCAACCCGCTGATGGGCTCATGTACTTGGACAGACTGGGAGAGAACCCCTAA	1500
Db	1441	ATCGACTGCAACCCGCTGATGGGCTCATGTACTTGGACAGACTGGGAGAGAACCCCTAA	1500
Qy	1501	ATCGAGTGTGCCAATTTGGATGGGACGAGCGGCTGTGTCATATGCTCTCCTCGGG	1560
Db	1501	ATCGAGTGTGCCAATTTGGATGGGACGAGCGGCTGTGTCATATGCTCTCCTCGGG	1560
Qy	1561	TGGCCCAACGGCTGGCCCTGGAACCTTCAAGAGGGGAAGCTCTACTGGGAGAGCGCAAG	1620
Db	1561	TGGCCCAACGGCTGGCCCTGGAACCTTCAAGAGGGGAAGCTCTACTGGGAGAGCGCAAG	1620
Qy	1621	ACAGACAAGATCGAGGTGATCAATGTTTGTATGGGACGAAGAGGGGGACCCCTCTCGGAGGAC	1680
Db	1621	ACAGACAAGATCGAGGTGATCAATGTTTGTATGGGACGAAGAGGGGGACCCCTCTCGGAGGAC	1680
Qy	1681	AAGCTCCGCACATTTTCGGGTTCACGCTGCTGGGGACATTCACTACTTGAATGACTGG	1740
Db	1681	AAGCTCCGCACATTTTCGGGTTCACGCTGCTGGGGACATTCACTACTTGAATGACTGG	1740
Qy	1741	CAGCGCGCAGCATCGAGCGGGTGCAACAGGTCAAGCGCAGCGGACGTCATATTGAC	1800
Db	1741	CAGCGCGCAGCATCGAGCGGGTGCAACAGGTCAAGCGCAGCGGACGTCATATTGAC	1800
Qy	1801	CAGCTGCCGACCTGATGGGGCTCAAAAGCTGTGAATGTGGCCAAAGGTCTGCGGAAACCAAC	1860

[illegible]

QY 1381 GCGCTCAACGCGCACCTCCGCAAGATCTGTGTGCGAGGACCTGAGAGCCCGAGCC 1440
DB 1400 GCGCTCAACGCGCACCTCCGCAAGATCTGTGTGCGAGGACCTGAGAGCCCGAGCC 1459
QY 1441 ATGCACTGCAACCCGCTGATGGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAAA 1500
DB 1460 ATGCACTGCAACCCGCTGATGGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAAA 1519
QY 1501 ATGAGTGTGCAACTTGGATGGGAGAGAGCGGCGTGTGCTGCTCAATGCGCTCCCTCGGG 1560
DB 1520 ATGAGTGTGCAACTTGGATGGGAGAGAGCGGCGTGTGCTGCTCAATGCGCTCCCTCGGG 1579
QY 1561 TGGCCCAACGCGCTCGGCTCGACCTGCAAGAGGGAAGCTTACTTGGGAGAGCCCAAG 1620
DB 1580 TGGCCCAACGCGCTCGGCTCGACCTGCAAGAGGGAAGCTTACTTGGGAGAGCCCAAG 1639
QY 1621 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAAGGCGGACCTCTCTGGAGGAC 1680
DB 1640 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAAGGCGGACCTCTCTGGAGGAC 1699
QY 1681 AAGCTCCGCAATTTTCGGGTTACGGCTTCAGGCTGTGGGGGACTTCATCTACTGGACTGACTGG 1740
DB 1700 AAGCTCCGCAATTTTCGGGTTTCACGCTGTGGGGGACTTCATCTACTGGACTGACTGG 1759
QY 1741 CAGCGCCGACGATCGAGCGGTGCAAAAGGTCAAGGTCAGCGGACGTCATCATTTGAC 1800
DB 1760 CAGCGCCGACGATCGAGCGGTGCAAAAGGTCAAGGTCAGCGGACGTCATCATTTGAC 1819
QY 1801 CAGCTGCCCGACCTGATGGGCTCAAAGCTGTGAATGTGGCCAAAGGTCTCGGAACCAAC 1860
DB 1820 CAGCTGCCCGACCTGATGGGCTCAAAGCTGTGAATGTGGCCAAAGGTCTCGGAACCAAC 1879
QY 1861 CCGTGTGGGACAGAAACGGGGGTGCAACCACTGTGCTTTCACACCCCAACCAAC 1920
DB 1880 CCGTGTGGGACAGAAACGGGGGTGCAACCACTGTGCTTTCACACCCCAACCAAC 1939
QY 1921 CCGTGTGGCTGCCCATCGGCTGAGCTGTGAGTGAATGATGAAGACCTGATCGTGCT 1980
DB 1940 CCGTGTGGCTGCCCATCGGCTGAGCTGTGAGTGAATGATGAAGACCTGATCGTGCT 1999
QY 1981 GAGGCTTCTTGGTCTTCAACAGCAGAGCGCGCATCCACAGGATCTCCCTCGAGACCAAT 2040
DB 2000 GAGGCTTCTTGGTCTTCAACAGCAGAGCGCGCATCCACAGGATCTCCCTCGAGACCAAT 2059
QY 2041 AACACGAGTGGCCATCCCGCTCAGCGGCTCAAGAGGCTCAGCGCTGAGCTTGCAT 2100
DB 2060 AACACGAGTGGCCATCCCGCTCAGCGGCTCAAGAGGCTCAGCGCTGAGCTTGCAT 2119
QY 2101 GTGTCCAAACCAACATCTACTGGACAGACGTCAGCTGAAGACCATCAGCGCGCTTC 2160
DB 2120 GTGTCCAAACCAACATCTACTGGACAGACGTCAGCTGAAGACCATCAGCGCGCTTC 2179
QY 2161 ATGAACGGGAGTCCGTGGAGCACTGTGTGAGTGTGGCTTGAATACCCGAGGGCATG 2220
DB 2180 ATGAACGGGAGTCCGTGGAGCACTGTGTGAGTGTGGCTTGAATACCCGAGGGCATG 2239
QY 2221 GCGTGTGATGGGCAAGAACTTACTTGGGCGGACACTGGGACCAACAGATCGAA 2280
DB 2240 GCGTGTGATGGGCAAGAACTTACTTGGGCGGACACTGGGACCAACAGATCGAA 2299
QY 2281 GTGGCGGCTCGACGGGCAAGTTCCGGCAAGTCTCGTGTGGAGGACTTGGACAAACCG 2340
DB 2300 GTGGCGGCTCGACGGGCAAGTTCCGGCAAGTCTCGTGTGGAGGACTTGGACAAACCG 2359
QY 2341 AGGTGCTGGCTCGATCCCAACGAAGGCTACATCTACTGACCGGAGTGGGGCGGCAAG 2400
DB 2360 AGGTGCTGGCTCGATCCCAACGAAGGCTACATCTACTGACCGGAGTGGGGCGGCAAG 2419
QY 2401 CCGAGGATGTCGGGCTTTCATGACGGGACCAACTGATCACTGCTGTGGACAGGTG 2460
DB 2420 CCGAGGATGTCGGGCTTTCATGACGGGACCAACTGATCACTGCTGTGGACAGGTG 2479
QY 2461 GCGCGGGCCAAACGACCTCAACCATTTGACTACGCTGACACGCGCTCTACTGGACCGACCTG 2520

DB 2480 GCGCGGGCCAAACGACCTCACCATTTGACTAGCTGACAGCGCTCTACTGGACCGACCTG 2539
QY 2521 GACACCAACATGATCGAGTGTGTCCAAACATGCTGGGTGAGGACGGGTGCTGATTTGCCGAC 2580
DB 2540 GACACCAACATGATCGAGTGTGTCCAAACATGCTGGGTGAGGACGGGTGCTGATTTGCCGAC 2599
QY 2581 GATCTCCGCGACCCGCTTCTGACGAGTGTGACGAGTGTATATATCTACTGACAGACTGG 2640
DB 2600 GATCTCCGCGACCCGCTTCTGAGTGTGACGAGTGTATATATCTACTGACAGACTGG 2659
QY 2641 AATCTGACAGCAATTTGAGCGGGCCGACAAGACTAGCGGCGGAAACCGCACCTCATCAG 2700
DB 2660 AATCTGACAGCAATTTGAGCGGGCCGACAAGACTAGCGGCGGAAACCGCACCTCATCAG 2719
QY 2701 GGCACCTTGTGATTTGATGGAATCTCTGTGTTCATCTCTCCCGCAGGATGGCTTC 2760
DB 2720 GGCACCTTGTGATTTGATGGAATCTCTGTGTTCATCTCTCCCGCAGGATGGCTTC 2779
QY 2761 AATGACTGTATGACAAACAGCGGCGAGTGTGGGACGCTGTGCTTGCCTCCCGGCGGC 2820
DB 2780 AATGACTGTATGACAAACAGCGGCGAGTGTGGGACGCTGTGCTTGCCTCCCGGCGGC 2839
QY 2821 CACCGCTGCGGCTGCGCTCACACTACACCTTGAACCCAGCAGCGGCAACTGACGCGC 2880
DB 2840 CACCGCTGCGGCTGCGCTCACACTACACCTTGAACCCAGCAGCGGCAACTGACGCGC 2899
QY 2881 CCGACCACTTCTTGTGCTTTCAGCCAGAAATCTGCCATCAGTCGATGATCCCGGACGAC 2940
DB 2900 CCGACCACTTCTTGTGCTTTCAGCCAGAAATCTGCCATCAGTCGATGATCCCGGACGAC 2959
QY 2941 CAGCAGCGCGGATCTCATCTGCCCTGATGAGCTGAGGAAACGTCAGGACCAACGATCGAC 3000
DB 2960 CAGCAGCGCGGATCTCATCTGCCCTGATGAGCTGAGGAAACGTCAGGACCAACGATCGAC 3019
QY 3001 TATGACCACTGGAACAAAGTTTCATCTACTGGTGTGATGGGCGGCAAAATCAAGCGAGCC 3060
DB 3020 TATGACCACTGGAACAAAGTTTCATCTACTGGTGTGATGGGCGGCAAAATCAAGCGAGCC 3079
QY 3061 AAGGACGAGGAGCCAGCCCTTGTGCTGAGCTCTGAGCGGAGGCGCAAAACCGACAC 3120
DB 3080 AAGGACGAGGAGCCAGCCCTTGTGCTGAGCTCTGAGCGGAGGCGCAAAACCGACAC 3139
QY 3121 AGGACGCGCCACGACCTCAGCATCGACATCTACAGCGCGGACACTGTTCTGAGACGTGCGAG 3180
DB 3140 AGGACGCGCCACGACCTCAGCATCGACATCTACAGCGCGGACACTGTTCTGAGACGTGCGAG 3199
QY 3181 GCCACCAATACCATCAACGTCACAGGTCGAGCGGGGAAGCCATGGGGGTGTTGCTGCT 3240
DB 3200 GCCACCAATACCATCAACGTCGAGGTCGAGCGGGGAAGCCATGGGGGTGTTGCTGCT 3259
QY 3241 GGGGACCGGACCAAGCCCGAGGCGCATCGTCTCAACGCGGAGCGGCTGCTGCTGCTTC 3300
DB 3260 GGGGACCGGACCAAGCCCGAGGCGCATCGTCTCAACGCGGAGCGGCTGCTGCTGCTTC 3319
QY 3301 ACCAATGACGAGGACCGGGGACGCAAGATCGAAACGCGCAGCGCTGGAACCGGACGCGC 3360
DB 3320 ACCAATGACGAGGACCGGGGACGCAAGATCGAAACGCGCAGCGCTGGAACCGGACGCGC 3379
QY 3361 GAGGTCTCTTCAACACCGGCTCATTCGCGCTGTGGCCCTGTGGTGGGACCAACACTG 3420
DB 3380 GAGGTCTCTTCAACACCGGCTCATTCGCGCTGTGGCCCTGTGGTGGGACCAACACTG 3439
QY 3421 GCGAGCTGTTCTGGGTGAGCGGACCTGAAGCGCATTTGAGAGCTGTGAGCTGTGAGG 3480
DB 3440 GCGAGCTGTTCTGGGTGAGCGGACCTGAAGCGCATTTGAGAGCTGTGAGCTGTGAGG 3499
QY 3481 GCGAACCGCTGACCTCGGAGGACGCGCAACATCTGTGAGCTCTGCGGCTTGACCATCTT 3540
DB 3500 GCGAACCGCTGACCTCGGAGGACGCGCAACATCTGTGAGCTCTGCGGCTTGACCATCTT 3559
QY 3541 GCGAACGATCTCTACTGTGATCGACCGCGCAGCAGATGATCGAGCTGTGAGGAGAC 3600

Db 3560 GCGAAGCATCTCTACTGGATCGACCGCCAGCAGCAGATGATCGAGCGTGTGAGAGAGACC 3619
Qy 3601 ACCGGGGCAAGCGAGCTCGCATCCAGGCGCGTGTGCGCCACCTCACTGGGATCCATGCA 3660
Db 3620 ACCGGGGCAAGCGAGCTCGCATCCAGGCGCGTGTGCGCCACCTCACTGGGATCCATGCA 3679
Qy 3661 GTGGAGGAGTCAAGCTCGAGGAGTTCCTCAGCCCAACCATGTGCCCGTGACAAATGGTGGC 3720
Db 3680 GTGGAGGAGTCAAGCTCGAGGAGTTCCTCAGCCCAACCATGTGCCCGTGACAAATGGTGGC 3739
Qy 3721 TGCTCCCAATCTGTATTCGCAAGGGTATGGGACACCAAGTGTCTATGCCAGTCCAC 3780
Db 3740 TGCTCCCAATCTGTATTCGCAAGGGTATGGGACACCAAGTGTCTATGCCAGTCCAC 3799
Qy 3781 CTCTGCTCTCTGCGAAGACTGTCTGACCTGTGGAGAGCGCCACCTGTCTCCCGGACCAAG 3840
Db 3800 CTCTGCTCTCTGCGAAGACTGTCTGACCTGTGGAGAGCGCCACCTGTCTCCCGGACCAAG 3859
Qy 3841 TTGTGATGTGCCACAGGGAGATCGACTGTATTCCTCCGGGGCTGTGGCGTGTGACGGCTTT 3900
Db 3860 TTGTGATGTGCCACAGGGAGATCGACTGTATTCCTCCGGGGCTGTGGCGTGTGACGGCTTT 3919
Qy 3901 CCCGAGTGGATGACACAGAGCGAGAGGCGTCCCGCGTGTGCTCCCGCGCCAGTTC 3960
Db 3920 CCCGAGTGGATGACACAGAGCGAGAGGCGTCCCGCGTGTGCTCCCGCGCCAGTTC 3979
Qy 3961 CCCTGCGCGCGGGTCACTGTGTGACCTGTGCGCTGCGCTGCGAGCGGAGGAGAGTGT 4020
Db 3980 CCCTGCGCGCGGGTCACTGTGTGACCTGTGCGCTGCGCTGCGAGCGGAGGAGAGTGT 4039
Qy 4021 CAGGACCGCTCAGAGCGAGCGAGCTGTGACGCGCATCTGCTGCGCCCAACAGTTCGGTGT 4080
Db 4040 CAGGACCGCTCAGAGCGAGCGAGCTGTGACGCGCATCTGCTGCGCCCAACAGTTCGGTGT 4099
Qy 4081 CGGAGCGCGAGTGTCTCATCAAACAGAGTGTGACTCTCTCCCGGAGCTGTATCGAC 4140
Db 4100 CGGAGCGCGAGTGTCTCATCAAACAGAGTGTGACTCTCTCCCGGAGCTGTATCGAC 4159
Qy 4141 GGCTCCGAGCAGCTCATGTGTGAATACCAAGCGCGCTCAGAGCAGCAGCGCGCCAC 4200
Db 4160 GGCTCCGAGCAGCTCATGTGTGAATACCAAGCGCGCTCAGAGCAGCAGCGCGCCAC 4219
Qy 4201 AGCAGTGCATCGGCGCGCTCATTTGGCATCATCTCTCTCTCTGCTATGGGTGTGTC 4260
Db 4220 AGCAGTGCATCGGCGCGCTCATTTGGCATCATCTCTCTCTCTGCTATGGGTGTGTC 4279
Qy 4261 TATTTTGTGTCAGCGCGTGTGTGCGAGCGTATGCGGGGCGCAACGCGGCGCTTCGCG 4320
Db 4280 TATTTTGTGTCAGCGCGTGTGTGCGAGCGTATGCGGGGCGCAACGCGGCGCTTCGCG 4339
Qy 4321 CACGAGTATGTCAGCGGAGCCCGCAGTGCCTCAATTTTCATAGCCCGGGCGGTTC 4380
Db 4340 CACGAGTATGTCAGCGGAGCCCGCAGTGCCTCAATTTTCATAGCCCGGGCGGTTC 4399
Qy 4381 CAGCATGCCCCCTTTCACAGGATCGCATGCGGAAAGTCCATGATGAGTCCCGTGAAGCTG 4440
Db 4400 CAGCATGCCCCCTTTCACAGGATCGCATGCGGAAAGTCCATGATGAGTCCCGTGAAGCTG 4459
Qy 4441 ATGGGGGGCGGGGGGGTGCCTCTGTACACCGGAACACGTCACAGGGGCTCGTTC 4500
Db 4460 ATGGGGGGCGGGGGGGTGCCTCTGTACACCGGAACACGTCACAGGGGCTCGTTC 4519
Qy 4501 AGCAGTCTGTCACAGCAGGAGCGCAGTGTACCGCGGATCTGAAACCGCGCGCTTC 4560
Db 4520 AGCAGTCTGTCACAGCAGGAGCGCAGTGTACCGCGGATCTGAAACCGCGCGCTTC 4579
Qy 4561 CCGGCCAGGACCCCTCCCTGTACAAATGACATGTTCTACTCTTCAAAATTCGCGCC 4620
Db 4580 CCGGCCAGGACCCCTCCCTGTACAAATGACATGTTCTACTCTTCAAAATTCGCGCC 4639
Qy 4621 ACTGCGAGACCGTACAGGCCCTATCATATTCAGGAATGGCGCCCGACGACGCCCTTCG 4680
Db 4640 ACTGCGAGACCGTACAGGCCCTATCATATTCAGGAATGGCGCCCGACGACGCCCTTCG 4699

Qy 4681 AGCACCGAGCTGTGTGACAGCGACTACAGCGCCAGCGCTGGAAGCCAGCAAGTACTAC 4740
Db 4700 AGCACCGAGCTGTGTGACAGCGACTACAGCGCCAGCGCTGGAAGCCAGCAAGTACTAC 4759
Qy 4741 CTGGATTTGAACTCGGACTCGAGACCCCTATCCACCCCGCCACCGAGAGGAGTCTTCCATCTC 4800
Db 4760 CTGGATTTGAACTCGGACTCGAGACCCCTATCCACCCCGCCACCGAGAGGAGTCTTCCATCTC 4819
Qy 4801 CTGTGCGGAGGAGCAGTGTGCCCGCTCGCGCGCCACCGAGAGGAGTCTTCCATCTC 4860
Db 4820 CTGTGCGGAGGAGCAGTGTGCCCGCTCGCGCGCCACCGAGAGGAGTCTTCCATCTC 4879
Qy 4861 TTCCGCGCCCTCCGTCCTCCCTCGACGAGTCTATCTGACCTCGCGCGGCGACCTGCGC 4920
Db 4880 TTCCGCGCCCTCCGTCCTCCCTCGACGAGTCTATCTGACCTCGCGCGGCGACCTGCGC 4939
Qy 4921 TTCTCTGTGCGCCCTGTAATAGTTTTTAAATATGAACAAAGAAAAAATATATTTTATGAT 4980
Db 4940 TTCTCTGTGCGCCCTGTAATAGTTTTTAAATATGAACAAAGAAAAAATATATTTTATGAT 4999
Qy 4981 TTAAAAATAAATAATATTGGGATTTTAAAAACATGAGAAATGTGAATCTGTGATGGGTG 5040
Db 5000 TTAATAATAAATAATATTGGGATTTTAAAAACATGAGAAATGTGAATCTGTGATGGGTG 5059
Qy 5041 GGCAGGGCTGGGAGAACTTTTGTA 5063
Db 5060 GGCAGGGCTGGGAGAACTTTTGTA 5082

RESULT 6
ACC45358
ID ACC45358 standard; DNA; 5120 BP.
XX AC ACC45358;
XX DT 02-JUN-2003 (first entry)
XX DE Human wild-type LRP5 gene.
XX KW Human; high bone mass; HBM; LRP5; LRP6; transgenic; bone mass modulation;
KW gene therapy; bone density modulation; bone strength; trabecular number;
KW bone size; bone tissue connectivity; bone disease; osteoporosis;
KW osteomalacia; rickets; Paget's disease; neoplasm of the bone; gene; ds.
XX OS Homo sapiens.
XX PN WO200292764-A2.
XX PD 21-NOV-2002.
XX PF 13-MAY-2002; 2002WO-US014876.
XX PR 11-MAY-2001; 2001US-0290071P.
PR 17-MAY-2001; 2001US-0291311P.
PR 01-FEB-2002; 2002US-0353058P.
PR 04-MAR-2002; 2002US-0361293P.
XX (GENO-) GENOME THERAPEUTICS CORP.
PA (AMHP) WYETH.
PI Babij P, Bex FJ, Yaworsky FJ, Bodine PV;
XX WPI; 2003-129278/12.
DR P-PSDB; ABR41093.
XX New transgenic animals (e.g. mice), useful as models for studying bone
PT density modulation, developing drugs for treating or preventing bone
PT diseases (e.g. osteoporosis), or diagnosing diseases characterized by
PT reduced bone density.
PS Disclosure; Fig 6; 603pp; English.
XX

1801 CAGCTGCCGACCTGATGGGGCTCAAAGCTGTGAATGTGGCCAAAGTCTGTGGAAACCAAC 1860
1820 CAGCTGCCGACCTGATGGGGCTCAAAGCTGTGAATGTGGCCAAAGTCTGTGGAAACCAAC 1879
1861 CCGTGTGGGACAGAAACCGGGGGTGCACCACTGTGCTTTTCAACCCCAACCAAC 1920
1880 CCGTGTGGGACAGAAACCGGGGGTGCACCACTGTGCTTTTCAACCCCAACCAAC 1939
1921 CGGTGTGGCTGCCCATCGGCTGTGAGCTGTGAGTGAATGAAGACCTGTGATGTGCTT 1980
1940 CGGTGTGGCTGCCCATCGGCTGTGAGCTGTGAGTGAATGAAGACCTGTGATGTGCTT 1999
1981 GAGGCTTCTTGGTCTTCAACAGACAGAGCGGCATCCACAGGATCTCCCTCGAGACCAAT 2040
2000 GAGGCTTCTTGGTCTTCAACAGACAGAGCGGCATCCACAGGATCTCCCTCGAGACCAAT 2059
2041 AACAAAGAGCTGGGCTCCCGCTCAACAGGCGTCAAGGAGGCTCAGCCCTGGACTTTGAT 2100
2060 AACAAAGAGCTGGGCTCCCGCTCAACAGGCGTCAAGGAGGCTCAGCCCTGGACTTTGAT 2119
2101 GTGTCCAAACCAACATCTACTGGACAGAGCTGAGCTGAAGACCATCAGCGGCGCTTC 2160
2120 GTGTCCAAACCAACATCTACTGGACAGAGCTGAGCTGAAGACCATCAGCGGCGCTTC 2179
2161 ATGAACGGGAGCTCGGTGGAGCAGCTGTGGAGTTTGGCTTGAATACCCCGAGGGCATG 2220
2180 ATGAACGGGAGCTCGGTGGAGCAGCTGTGGAGTTTGGCTTGAATACCCCGAGGGCATG 2239
2221 GCCGTGTGACTGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGAACTCGAA 2280
2240 GCCGTGTGACTGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGAACTCGAA 2299
2281 GTGGGCGGCTGGAGCGGCGAGTTCCGGCAAGTCTCTGCTGTGGAGGACTTGGACAAACCG 2340
2300 GTGGGCGGCTGGAGCGGCGAGTTCCGGCAAGTCTCTGCTGTGGAGGAGACTTGGACAAACCG 2359
2341 AGGTGCTGGCGCTGGATCCCAACAAAGGCTACATCTACTGGACCGAGTGGGGGGCAAG 2400
2360 AGGTGCTGGCGCTGGATCCCAACAAAGGCTACATCTACTGGACCGAGTGGGGGGCAAG 2419
2401 CCGAGGATGTGCGGGCTTTCATGGACGGGACCAACTGCTGATGACGCTGTGGTGGACAAAGTG 2460
2420 CCGAGGATGTGCGGGCTTTCATGGACGGGACCAACTGCTGATGACGCTGTGGTGGACAAAGTG 2479
2461 GCGCGGGCAACGACCTCAACATTTGACTACGCTGACAGCGCTCTACTGACCGACCTG 2520
2480 GCGCGGGCAACGACCTCAACATTTGACTACGCTGACAGCGCTCTACTGACCGACCTG 2539
2521 GACACCAACATGATCGAGTCTGCAACATGCTGGGTCAAGGAGCGGTCTGATTTGCGGAC 2580
2540 GACACCAACATGATCGAGTCTGCAACATGCTGGGTCAAGGAGCGGTCTGATTTGCGGAC 2599
2581 GATCTCCGCGACCGCTTCCGCTGTGACGAGTACAGGATTTATCTACTGGACAGACTGG 2640
2600 GATCTCCGCGACCGCTTCCGCTGTGACGAGTACAGGATTTATCTACTGGACAGACTGG 2659
2641 AATCTGACACGATTTGAGCGGGCGGACAAAGATAGCGGCGGGACCGACCCCTCATCCAG 2700
2660 AATCTGACACGATTTGAGCGGGCGGACAAAGATAGCGGCGGGACCGACCCCTCATCCAG 2719
2701 GGCACCTGGAGCTTCTGATGAGATCTCTGTGTGTTTCACTCTCCCGCAGGATGGCTTC 2760
2720 GGCACCTGGAGCTTCTGATGAGATCTCTGTGTGTTTCACTCTCCCGCAGGATGGCTTC 2779
2761 AATGATGTATGACAAACAAAGGCGAGTGTGGGAGCTGTGCTTTGCCATCCCGGGGCGC 2820
2780 AATGATGTATGACAAACAAAGGCGAGTGTGGGAGCTGTGCTTTGCCATCCCGGGGCGC 2839
2821 CACGCTCGGCTGGCTCTCACTACACCTTGACCCGACGAGCGGCAATGTGACGCGCG 2880
2840 CACGCTCGGCTGGCTCTCACTACACCTTGACCCGACGAGCGGCAATGTGACGCGCG 2899

2881 CCCACCACCTTCTTGTCTTTCAGCCAGAAATCTGCCATCAGTCGATGATCCCGACGAC 2940
2900 CCCACCACCTTCTTGTCTTTCAGCCAGAAATCTGCCATCAGTCGATGATCCCGACGAC 2959
2941 CAGCACAGCCCGGATCTCATCTCTGCTGATGATGAGGAACTGAGGAACTGAGGAACTGAG 3000
2960 CAGCACAGCCCGGATCTCATCTCTGCTGATGATGAGGAACTGAGGAACTGAGGAACTGAG 3019
3001 TATGACCCACTGGACAGATTTCTACTTGGTGGATGGGCGCCAGAACTCAGCGAGCC 3060
3020 TATGACCCACTGGACAGATTTCTACTTGGTGGATGGGCGCCAGAACTCAGCGAGCC 3079
3061 AAGGACGAGCGGACCCAGCCCTTTTGTGTGACCTCTCTGAGCCAGGCGCAAAACCCAGAC 3120
3080 AAGGACGAGCGGACCCAGCCCTTTTGTGTGACCTCTCTGAGCCAGGCGCAAAACCCAGAC 3139
3121 AGGACGCGCCACGACCTCAGCATCTGACATCTACAGCGGACACTTGTCTGAGCTGCGAG 3180
3140 AGGACGCGCCACGACCTCAGCATCTGACATCTACAGCGGACACTTGTCTGAGCTGCGAG 3199
3181 GCCACCAATACCATCAAGTCCACAGGCTGACGCGGGAGGCATGGGGGTGTGCTGCT 3240
3200 GCCACCAATACCATCAAGTCCACAGGCTGACGCGGGAGGCATGGGGGTGTGCTGCT 3259
3241 GGGGACCGCGACAAAGCCAGGGCCATCGTCTCAACGCGGAGCGAGGCTACCTGTACTTC 3300
3260 GGGGACCGCGACAAAGCCAGGGCCATCGTCTCAACGCGGAGCGAGGCTACCTGTACTTC 3319
3301 ACCAACATGACGAGCCGCGGACGCAAGATCGAACGCGCAGCCCTGAGACCGGACCGAGCGC 3360
3320 ACCAACATGACGAGCCGCGGACGCAAGATCGAACGCGCAGCCCTGAGACCGGACCGAGCGC 3379
3361 GAGGTCTCTTTCACACCGGCTCATCCGCCCTGTGGCCCTGTGGTGGACAAACACTG 3420
3380 GAGGTCTCTTTCACACCGGCTCATCCGCCCTGTGGCCCTGTGGTGGACAAACACTG 3439
3421 GCGAAGCTGTCTTGGGTGGACCGGACCTGAAGCGCATTTAGAGCTGTGACCTGTCAAGG 3480
3440 GCGAAGCTGTCTTGGGTGGACCGGACCTGAAGCGCATTTAGAGCTGTGACCTGTCAAGG 3499
3481 GCCAACCGCTGACCTCTGGAGGACCGCAACATCTGTGACGCTCTTGGGCTGACCACTCTT 3540
3500 GCCAACCGCTGACCTCTGGAGGACCGCAACATCTGTGACGCTCTTGGGCTGACCACTCTT 3559
3541 GCGAAGCATCTTACTGATCGACCGGCGGACGAGATGATCGAGCGGTGTGGAGGAGAC 3600
3560 GCGAAGCATCTTACTGATCGACCGGCGGACGAGATGATCGAGCGGTGTGGAGGAGAC 3619
3601 ACCGGGACAAAGCGGACCTCGCATCCAGGGCGGTGTGGCCCACTCAGCTGGCATCCATGCA 3660
3620 ACCGGGACAAAGCGGACCTCGCATCCAGGGCGGTGTGGCCCACTCAGCTGGCATCCATGCA 3679
3661 GTGAGGAAAGTACGCTCTGGAGGATTTCTAGCCCAACCATGTGCCCGTGAACAATGGTGGC 3720
3680 GTGAGGAAAGTACGCTCTGGAGGATTTCTAGCCCAACCATGTGCCCGTGAACAATGGTGGC 3739
3721 TGCTCCACATCTGTATTTGCCAAGGTGATGGGACACCAAGTGTCTATGCGGCTGAC 3780
3740 TGCTCCACATCTGTATTTGCCAAGGTGATGGGACACCAAGTGTCTATGCGGCTGAC 3799
3781 CTGCTGCTCTGCAAGACCTGCTGACCTGAGAGGCGGCGCACTGCTGCCCGGACGAG 3840
3800 CTGCTGCTCTGCAAGACCTGCTGACCTGAGAGGCGGCGCACTGCTGCCCGGACGAG 3859
3841 TTTGATGTGCAACAGGGGAGATGACTGTATCCCGGGGCTTGCGCTGTGACGGCTTT 3900
3860 TTTGATGTGCAACAGGGGAGATGACTGTATCCCGGGGCTTGCGCTGTGACGGCTTT 3919
3901 CCGGAGTGCATGACAGAGGACGAGGAGGCTGCGCCGCTGTGCTCGCGCGGCGGCTTC 3960
3920 CCGGAGTGCATGACAGAGGACGAGGAGGCTGCGCCGCTGTGCTCGCGCGGCGGCTTC 3979
3961 CCCTGCGCGGGGTCTAGTGTGGACCTGTGCGCTGCGCTGCGAGCGGCGAGGACGCTGT 4020

Db 200 CGGCTGGTGGACGCGCGGAGTCAAGCTGGAGTCCACCATCTGGTCTCAGCGGCTGGAG 259
Qy 241 GATCGGCGCAAGTGGACTCCAGTCTTTCGAAGGAGCGGTGTACTGGACAGAGTGGAC 300
Db 260 GATCGGCGCGAGTGGACTTCCAGTCTTTCGAAGGAGCGGTGTACTGGACAGAGTGGAC 319
Qy 301 GAGGAGGCAATCAAGCAGACCTACTGAAACAGACGCGGCGCGGTGTCAGAAAGTGGTC 360
Db 320 GAGGAGGCAATCAAGCAGACCTACTGAAACAGACGCGGCGCGGTGTCAGAAAGTGGTC 379
Qy 361 ATCTCCGCGCTGGTCTCTCCCGACGCGCTCGCTCGAGCTGGGTGGGCAAGAGTGTAC 420
Db 380 ATCTCCGCGCTGGTCTCTCCCGACGCGCTCGCTCGAGCTGGGTGGGCAAGAGTGTAC 439
Qy 421 TGGACGACTCAGAGACCAACCGCATCGAGTGGCCAACTCAATGGCAGACATCCCGGAAG 480
Db 440 TGGACGACTCAGAGACCAACCGCATCGAGTGGCCAACTCAATGGCAGACATCCCGGAAG 499
Qy 481 GTGCTCTTCTGGCAGGACCTTTGACAGCCTTAGGGCCATTCGCTTTGGACCCCGCTCACGGG 540
Db 500 GTGCTCTTCTGGCAGGACCTTTGACAGCCTTAGGGCCATTCGCTTTGGACCCCGCTCACGGG 559
Qy 541 TACATGTACTGGACAGACTGGGTGGAGCGCCCGGATTTGAGCGGGCAGGATGGATGGC 600
Db 560 TACATGTACTGGACAGACTGGGTGGAGCGCCCGGATTTGAGCGGGCAGGATGGATGGC 619
Qy 601 AGCACCGGAAGATCATTTGTGGACTCGGACATTTACTGGCCCAATGACTGACCATCGAC 660
Db 620 AGCACCGGAAGATCATTTGTGGACTCGGACATTTACTGGCCCAATGACTGACCATCGAC 679
Qy 661 CTGGAGGACGAGAAGCTCTACTGGGCTGACGCCAAGCTCAGCTTTCATCCACCGTGCACAC 720
Db 680 CTGGAGGACGAGAAGCTCTACTGGGCTGACGCCAAGCTCAGCTTTCATCCACCGTGCACAC 739
Qy 721 CTGACCGGCTCTCCGCGAGAGGTGGTGGAGGGCAGCTGACGACCCCTTCCGCGCTG 780
Db 740 CTGACCGGCTCTCCGCGAGAGGTGGTGGAGGGCAGCTGACGACCCCTTCCGCGCTG 799
Qy 781 ACGCTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCCATCCATGCCCTGC 840
Db 800 ACGCTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCCATCCATGCCCTGC 859
Qy 841 AACAGCGCACTGGGGGGAAGAGGAGATCTGTAGTGCCCTCTACTCACCCATGGAC 900
Db 860 AACAGCGCACTGGGGGGAAGAGGAGATCTGTAGTGCCCTCTACTCACCCATGGAC 919
Qy 901 ATCCAGGTGCTGACGACGAGCGGACGCTTTCTTCCACACTCGCTGTGAGGAGGACAAT 960
Db 920 ATCCAGGTGCTGACGACGAGCGGACGCTTTCTTCCACACTCGCTGTGAGGAGGACAAT 979
Qy 961 GCGGCTGCTCCACCTGTGCTCTGTCCCAAGCGAGCCTTTCTACATGCCCTGC 1020
Db 980 GCGGCTGCTCCACCTGTGCTCTGTCCCAAGCGAGCCTTTCTACATGCCCTGC 1039
Qy 1021 CCCACGGGTGTGACGTGCAAGCAACCGGACGACGCTGTAAAGGAGGAGCGGAGGAGTG 1080
Db 1040 CCCACGGGTGTGACGTGCAAGCAACCGGACGACGCTGTAAAGGAGGAGCGGAGGAGTG 1099
Qy 1081 CTGCTGCTGGCGCGGCGGACGACCTAGGAGGATCTCGCTGGACAGCGCGGACTTCACC 1140
Db 1100 CTGCTGCTGGCGCGGCGGACGACCTAGGAGGATCTCGCTGGACAGCGCGGACTTCACC 1159
Qy 1141 GACATCGTGTGAGGTGGACGACATCCGACGCCATTTGCCATCGACTAGACCCGCTA 1200
Db 1160 GACATCGTGTGAGGTGGACGACATCCGACGCCATTTGCCATCGACTAGACCCGCTA 1219
Qy 1201 GAGGGCTATGTCTACTGGACAGATGAGAGGTGGGGCCATCCGACGAGGGGTACTTGGAC 1260
Db 1220 GAGGGCTATGTCTACTGGACAGATGAGAGGTGGGGCCATCCGACGAGGGGTACTTGGAC 1279
Qy 1261 GGGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGCCGATGGCATCGCGGTG 1320
Db 1280 GGGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGCCGATGGCATCGCGGTG 1339

Qy 1321 GACTGGGTGGCCCGAABACCTCTACTGGACCGACACGGGCAAGGACCGCATCGAGGTGACG 1380
Db 1340 GACTGGGTGGCCCGAABACCTCTACTGGACCGACACGGGCAAGGACCGCATCGAGGTGACG 1399
Qy 1381 CGCCTCAACGGCAGCTCCCGCAAGATCCTGTGTTCGAGGACCTTGGAGAGCCCCGAGCC 1440
Db 1400 CGCCTCAACGGCAGCTCCCGCAAGATCCTGTGTTCGAGGACCTTGGAGAGCCCCGAGCC 1459
Qy 1441 ATCGCACTGCACCCCGTGTATGGGCTCATGTACTGGACAGACTGGGAGAGAACCCCTAAA 1500
Db 1460 ATCGCACTGCACCCCGTGTATGGGCTCATGTACTGGACAGACTGGGAGAGAACCCCTAAA 1519
Qy 1501 ATCGAGTGTGCCTAACTTTGGATGGGACGAGCGGTGTGTGTCTCAATGCCTCCCTCGG 1560
Db 1520 ATCGAGTGTGCCTAACTTTGGATGGGACGAGCGGTGTGTGTCTCAATGCCTCCCTCGG 1579
Qy 1561 TGGCCCAACGGCCTTGGCCCTTGGACCTCGAGAGGGGAACTCTTACTGGGGAGAGCCCAAG 1620
Db 1580 TGGCCCAACGGCCTTGGCCCTTGGACCTCGAGAGGGGAACTCTTACTGGGGAGAGCCCAAG 1639
Qy 1621 ACAGACAAGATCGAGGTGTATCAATGTGTATGGGACAGAGGGCGGACCCCTCTGGAGAC 1680
Db 1640 ACAGACAAGATCGAGGTGTATCAATGTGTATGGGACAGAGGGCGGACCCCTCTGGAGAC 1699
Qy 1681 AAGCTCCCGCACATTTTCGGGTTTCACTGTGGGGGACTTCTACTTCTGACTGACTGG 1740
Db 1700 AAGCTCCCGCACATTTTCGGGTTTCACTGTGGGGGACTTCTACTTCTGACTGACTGG 1759
Qy 1741 CAGCGCGCAGCATCGAGCGGGTGCAACAAGTCAAGGTCAGGCGGAGCGTCACTATTGAC 1800
Db 1760 CAGCGCGCAGCATCGAGCGGGTGCAACAAGTCAAGGTCAGGCGGAGCGTCACTATTGAC 1819
Qy 1801 CAGTGTCCGACACTGATGGGGCTCAAAAGCTGTGATGTGGCCAAAGTGTCTGGAAACCAAC 1860
Db 1820 CAGTGTCCGACACTGATGGGGCTCAAAAGCTGTGATGTGGCCAAAGTGTCTGGAAACCAAC 1879
Qy 1861 CCGTGTCCGACAGGAAACGGGGGTGAGCCACTGTGTCTTTCACACCCACGCAACC 1920
Db 1880 CCGTGTCCGACAGGAAACGGGGGTGAGCCACTGTGTCTTTCACACCCACGCAACC 1939
Qy 1921 CCGTGTGGCTGCCCATTCGCGCTTGGAGCTGTGATGACATGAAGACTTGCATCGTGCCT 1980
Db 1940 CCGTGTGGCTGCCCATTCGCGCTTGGAGCTGTGATGACATGAAGACTTGCATCGTGCCT 1999
Qy 1981 GAGGCTTCTTGTCTTTCACAGCAGAGCGCCATTCACAGGATCTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTTCTTGTGTCTTTCACAGCAGAGCGCCATTCACAGGATCTCCCTCGAGACCAAT 2059
Qy 2041 AACAAAGCAGCTGGCCATCCCGCTCACGGGCGTCAAGGAGGCTCAGGCCCTGGACTTTGAT 2100
Db 2060 AACAAAGCAGCTGGCCATCCCGCTCACGGGCGTCAAGGAGGCTCAGGCCCTGGACTTTGAT 2119
Qy 2101 GTGTCCAAACCAACATCTACTGGACAGAGCTCAGCCTGAAGACCATCAGCGCGGCTTC 2160
Db 2120 GTGTCCAAACCAACATCTACTGGACAGAGCTCAGCCTGAAGACCATCAGCGCGGCTTC 2179
Qy 2161 ATGAAACGGGAGCTCGGTGGAGCAGCTGTGGAGTTTGGCCCTTGTACTACCCCGAGGCAATG 2220
Db 2180 ATGAAACGGGAGCTCGGTGGAGCAGCTGTGGAGTTTGGCCCTTGTACTACCCCGAGGCAATG 2239
Qy 2221 GCCCTTGACTGGATGGCAAGAACTCTACTGGGCGCAGACTGGGACCAACAGAAATCGAA 2280
Db 2240 GCCCTTGACTGGATGGCAAGAACTCTACTGGGCGCAGACTGGGACCAACAGAAATCGAA 2299
Qy 2281 GTGGCGGCTGAGACGGGCAAGTTCCGCAAGTCTCTGTGTGGAGGAGTCTTGAACAAACCCG 2340
Db 2300 GTGGCGGCTGAGACGGGCAAGTTCCGCAAGTCTCTGTGTGGAGGAGTCTTGAACAAACCCG 2359
Qy 2341 AGTGTGCTGGCCCTTGGATCCCAACAGGGCTCATCTACTTGGACCGGAGTGGGGCGGCAAG 2400
Db 2360 AGTGTGCTGGCCCTTGGATCCCAACAGGGCTCATCTACTTGGACCGGAGTGGGGCGGCAAG 2419

QY	2401	CCGAGGATGCTGGCGGCTTCAATGACGGGACCAACTGCATGACGCTGGTGACAGGTTG	2460		Db	3500	GCCAAACCGCCCTGACCCCTGGAGGACGCCAAACATCGTGTGAGCCTCTGGGGCTGACCAATCCTT	3559	
Db	2420	CGAGGATCTGTGGGGCTTTCATGACGGGACCAACTGCATGACGCTGGTGACAAAGTTG	2479		QY	3541	GGCAAGCATCTCTACTGATCGAACCGCCAGCAGCAGATGATCGAGCGTGTGGAGAGACC	3600	
QY	2461	GGCCGGGCAACGACCTCACCATTTGACTACGTGACCGAGCGCTTACTTGGACCGACCTG	2520		Db	3560	GGCAAGCATCTCTACTGATCGAACCGCCAGCAGCAGATGATCGAGCGTGTGGAGAGACC	3619	
Db	2480	GGCCGGGCAACGACCTCACCATTTGACTACGTGACCGAGCGCTTACTTGGACCGACCTG	2539		QY	3601	ACCGGGGCAAGAGCGGACCTGCGATCAGGGCCGTGTGCCCACTCACTTGGCATCCATGCA	3660	
QY	2521	GACACCAATCATGATCGAGTCTGTCMAACATGTCTGGTCAAGGAGCGGTCGTGATGCCGAC	2580		Db	3620	ACCGGGGCAAGAGCGGACCTGCGATCCAGGGCGGTGTGCCCACTCACTTGGCATCCATGCA	3679	
Db	2540	GACACCAATCATGATCGAGTCTGTCMAACATGTCTGGTCAAGGAGCGGTCGTGATGCCGAC	2599		QY	3661	GTGGAGGAAGTCAGCTCGGAGGATTTCTCAGCCACCACTGTGCCGTGACAAATGGTGGC	3720	
QY	2581	GATCTCCCGCACCGCTCGGTCTGACGAGTACAGCGATTATATCTACTGACAGACTGG	2640		Db	3680	GTGGAGGAAGTCAGCTCGGAGGATTTCTCAGCCACCACTGTGCCGTGACAAATGGTGGC	3739	
Db	2600	GATCTCCCGCACCGCTCGGTCTGACGAGTACAGCGATTATATCTACTGACAGACTGG	2659		QY	3721	TGCTCCCAACATCTGTATTGGCCAGGCTGATGGGACACCAACAGTGTCTATGGCCAGTCCAC	3780	
QY	2641	AATCTGCACAGCATTTGAGCGGGCCGACAAAGACTAGCGGCCGGAAACCGCACCTCATCCAG	2700		Db	3740	TGCTCCCAACATCTGTATTGGCCAGGCTGATGGGACACCAACAGTGTCTATGGCCAGTCCAC	3799	
Db	2660	AATCTGCACAGCATTTGAGCGGGCCGACAAAGACTAGCGGCCGGAAACCGCACCTCATCCAG	2719		QY	3781	CTCGTGTCTCTGCAGAACCTGCTGACCTGTGGAGAGCGGCCCACTGTCTCCCGGACGAG	3840	
QY	2701	GGCCACTTGGATCTGATGAGCAATCTGTGTGTTCACCTCTCCCGCCAGATGGCCCTC	2760		Db	3800	CTCGTGTCTCTGCAGAACCTGCTGACCTGTGGAGAGCGGCCCACTGTCTCCCGGACGAG	3859	
Db	2720	GGCCACTTGGATCTGATGAGCAATCTGTGTGTTCACCTCTCCCGCCAGATGGCCCTC	2779		QY	3841	TTTGCATGTGTCCACAGGGGAGATCGACTGTATCCCGGGGCTGTGCGCTGTGACGGCTTT	3900	
QY	2761	AATGACTGTATGCACAAACGGGCGAGTGTGGCAGCTGTGCCATCCCGGGCGC	2820		Db	3860	TTTGCATGTGTCCACAGGGGAGATCGACTGTATCCCGGGGCTGTGCGCTGTGACGGCTTT	3919	
Db	2780	AATGACTGTATGCACAAACGGGCGAGTGTGGCAGCTGTGCCATCCCGGGCGC	2839		QY	3901	CCCGAGTGCATGACACAGAGCGAGAGGGCTGTCCCGCTGTGTCTCCCGCCCGCAGTTC	3960	
QY	2821	CACCGTCTGGCTGTGGCTCTACCTACACCTCTGACCCCGCAGCAGCGCAACTGACGCCG	2880		Db	3920	CCCGAGTGCATGACACAGAGCGAGAGGGCTGTCCCGCTGTGTCTCCCGCCCGCAGTTC	3979	
Db	2840	CACCGTCTGGCTGTGGCTCTACCTACACCTCTGACCCCGCAGCAGCGCAACTGACGCCG	2899		QY	3961	CCCTTGGCGGGGTCTAGTGTGTGGAACCTGTGGCCCTGTGCGTGTGGAAGGAGGAGACTGT	4020	
QY	2881	CCGACCACTTCTTGTCTTTCAGCCAGAAATCTGCGCATCTGCGGATGATCCCGGACGAC	2940		Db	3980	CCCTTGGCGGGGTCTAGTGTGTGGAACCTGTGGCCCTGTGCGTGTGGAAGGAGGAGACTGT	4039	
Db	2900	CCGACCACTTCTTGTCTTTCAGCCAGAAATCTGCGCATCTGCGGATGATCCCGGACGAC	2959		QY	4021	CAGGACCGCTCAGACGAGGGCGACTGTGACGCGCATCTGCGTCCCAACAGTTCGGGTGT	4080	
QY	2941	CAGCAGCGCCGGATCTCATCTGCGCTGTGATGAGACTGAGGAACTGCAAGGCCATGAC	3000		Db	4040	CAGGACCGCTCAGACGAGGTGGAATGTGACGCGCATCTGCGTCCCAACAGTTCGGGTGT	4099	
Db	2960	CAGCAGCGCCGGATCTCATCTGCGCTGTGATGAGACTGAGGAACTGCAAGGCCATGAC	3019		QY	4081	GCAGCGGCCAGTGTCTCTCATCAAAACAGCAGTGTGCACTCTTTCGCCGATGTATGAC	4140	
QY	3001	TATGACCACTTGCACAAAGTTTCATCTACTGGTGTGATGGGCCCGCAGAACATCAAGCGAC	3060		Db	4100	GCAGCGGCCAGTGTCTCTCATCAAAACAGCAGTGTGCACTCTTTCGCCGATGTATGAC	4159	
Db	3020	TATGACCACTTGCACAAAGTTTCATCTACTGGTGTGATGGGCCCGCAGAACATCAAGCGAC	3079		QY	4141	GGCTCCGACGAGCTCATGTGTGAAATCACAAGCGCGCTCAGACGACAGCGCGGCCAC	4200	
QY	3061	AAGGACGAGCGGACCCAGCCCTTTTGTGACCTCTCTGAGCCAGGCCAAACCCAGAC	3120		Db	4160	GGCTCCGACGAGCTCATGTGTGAAATCACAAGCGCGCTCAGACGACAGCGCGGCCAC	4219	
Db	3080	AAGGACGAGCGGACCCAGCCCTTTTGTGACCTCTCTGAGCCAGGCCAAACCCAGAC	3139		QY	4201	AGCAGTGCCATTCGGGCGCCGTCATTGGGCATCATCTCTCTCTTCTGTCATGGGTGGTGC	4260	
QY	3121	AGGACGCCCCACGACCTCAGCATCGACATCTACAGCCCGGACACTGTCTGAGACGTGCGAG	3180		Db	4220	AGCAGTGCCATTCGGGCGCCGTCATTGGGCATCATCTCTCTCTTCTGTCATGGGTGGTGC	4279	
Db	3140	AGGACGCCCCACGACCTCAGCATCGACATCTACAGCCCGGACACTGTCTGAGACGTGCGAG	3199		QY	4261	TATTTTGTGTGCCAGCGGTGTGTGCCAGCGCTATGCGGGGGCCAAACGGGCGCTTCCCG	4320	
QY	3181	GCCACCAATACATCAACGTCACAGGTCGAGCGGGGAAGCCATGGGGGTGGTCTGCGT	3240		Db	4280	TATTTTGTGTGCCAGCGGTGTGTGCCAGCGCTATGCGGGGGCCAAACGGGCGCTTCCCG	4339	
Db	3200	GCCACCAATACATCAACGTCACAGGTCGAGCGGGGAAGCCATGGGGGTGGTCTGCGT	3259		QY	4321	CACGAGTGTGACGCGGAGCCCGCAGTGCCTCTCAATTTTCATAGCCCGGGGGTTC	4380	
QY	3241	GGGACCGCGACAGCCCGGGCCATCGTCTGTCACCGGAGCGAGGGTACTCTGACTTC	3300		Db	4340	CACGAGTGTGACGCGGAGCCCGCGACGCTGCTCAATTTTCATAGCCCGGGGGTTC	4399	
Db	3260	GGGACCGCGACAGCCCGGGCCATCGTCTGTCACCGGAGCGAGGGTACTCTGACTTC	3319		QY	4381	CAGCATGGCCCTTTTCACAGGCATCGCATGGGAAAGTCCATGATGAGTCCGTGAGCGCTG	4440	
QY	3301	ACCAACATGACGAGCCCGGGACGCAAGATCGAAACGCGCAGCCCTGGAAGCGCACCGAGCGC	3360		Db	4400	CAGCATGGCCCTTTTCACAGGCATCGCATGGGAAAGTCCATGATGAGTCCGTGAGCGCTG	4459	
Db	3320	ACCAACATGACGAGCCCGGGACGCAAGATCGAAACGCGCAGCCCTGGAAGCGCACCGAGCGC	3379		QY	4441	ATGGGGGGCGGGGGGGTGCCTCTGACGACCGGAAACACAGTCAAGGGGGCTCTGTC	4500	
QY	3361	GAGTCTCTTTCACACCGGCTCATTCGCGCTGTGGCCCTGTGGTGTGGGAAACACACTG	3420		Db	4460	ATGGGGGGCGGGGGGGTGCCTCTTACGACCGGAAACACAGTCAAGGGGGCTCTGTC	4519	
Db	3380	GAGTCTCTTTCACACCGGCTCATTCGCGCTGTGGCCCTGTGGTGTGGGAAACACACTG	3439		QY	4501	AGCAGCTCTGTCAGACGAAAGGGCCAGCTGTACCGCGCGATCTCTGAAACCGCGCCCTCC	4560	
QY	3421	GGCAAGCTGTTCTGGTGTGACCGCGACTGTAAGCGCATTTGAGAGCTGTGACTCTCAGGG	3480		Db	4520	AGCAGCTCTGTCAGACGAAAGGGCCAGCTGTACCGCGCGATCTCTGAAACCGCGCCCTCC	4579	
Db	3440	GGCAAGCTGTTCTGGTGTGACCGCGACTGTAAGCGCATTTGAGAGCTGTGACTCTCAGGG	3499		QY	4561	CCGCGCACGAGCCCTCTCTGTACCAATGGACATGTTCTTACTCTTCAAAATTCGCGCC	4620	
QY	3481	GCCAAACCGCCTGACCCCTGGAGGAGCGCCAAACATCGTGTGAGCCTCTGGGGCTGACCATCCTT	3540						

4580	CGGCCACGAGACCCTCCCTGTGTACACATGGACATGTTCTTACTCTTCAAACATTCGGGCC	4639
4621	ACTGCGAGACCGTACAGGCCCTACATCATTTGAGGAATGGCGCCCCCGACGACGCCCTGCG	4680
4640	ACTGCGAGACCGTACAGGCCCTACATCATTTGAGGAATGGCGCCCCCGACGACGCCCTGCG	4699
4681	AGCACCGACGTGTGTACAGCGACTACAGCGCCAGCGCTGGAGGCGCCAGCAGTACTAC	4740
4700	AGCACCGACGTGTGTACAGCGACTACAGCGCCAGCGCTGGAGGCGCCAGCAGTACTAC	4759
4741	CTGGATTTGAACCTGGACTCAGACCCCTATCCACCCCCACCCACCGCCCCACAGCCAGTAC	4800
4760	CTGGATTTGAACCTGGACTCAGACCCCTATCCACCCCCACCCACCGCCCCACAGCCAGTAC	4819
4801	CTGTGCGCGGAGGACAGCTGCCGCCCTCGCCGCCACCGAGGAGGAGTACTTCCATCTC	4860
4820	CTGTGCGCGGAGGACAGCTGCCGCCCTCGCCGCCACCGAGGAGGAGTACTTCCATCTC	4879
4861	TTCCGCGCCCTCCGTCCTCCCTTGACAGGAGTCACTCTGACCTCGCCCGGCGCCACTCTGGC	4920
4880	TTCCGCGCCCTCCGTCCTCCCTTGACAGGAGTCACTCTGACCTCGCCCGGCGCCACTCTGGC	4939
4921	TTCTCTGTGCGCCCTGTAAATAGTTTTAAATATGAAACAAAGAAAAATATATTTATGAT	4980
4940	TTCTCTGTGCGCCCTGTAAATAGTTTTAAATATGAAACAAAGAAAAATATATTTATGAT	4999
4981	TTAAAAATAAATAATATCGGATTTTAAAAACATGAGAAATGTGAACCTGTGATGGGGTG	5040
5000	TTAAAAATAAATAATATCGGATTTTAAAAACATGAGAAATGTGAACCTGTGATGGGGTG	5059
5041	GGCAGGCGCTGGGAGAACTTTGTA	5063
5060	GGCAGGCGCTGGGAGAACTTTGTA	5082

RESULT 8

ADE82639
ID ADE82639 standard; DNA; 5120 BP.

AC ADE82639;

29-JAN-2004 (first entry)

Human Zmax1 gene encoding sequence.

LRP5; LRP6; HBM; Dkk activity; Osteopathic; Antiinflammatory;
 Antiarthritic; bone mass disorders; osteoporosis; hypercalcaemia;
 hyperostosis; osteogenesis; Wnt signaling; ds.

XX
OS
Homo sapiens.

	Key	Location/Qualifiers
AA	CDS	71. .4915
FH		
FT		
FT		/*tag= a

PN WO200292015-A2.

21-NOV-2002.

17-MAY-2002: 2002WO-US015982.

XX
PB 17-MAY-2001. 2001US-0291311P

PR I/-MAY-2001; 2001US-0291311P.
PR 01-FEB-2002; 2002US-0353058P

PR 01-FEB-2002; 2002US-0353058F.
PR 04-MAR-2002; 2002US-0361293P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.
PA (AMHP) WYETH.

XX
PI Allen K, Anisowicz A, Bhat BM, Damagnez V, Robinson JA;
PI Yaworsky PJ;

XX WPI; 2003-129219/12.
DR P-PSDB; ADE82427.
DR

Db 680 CTGGAGGACGAGAAGCTCTACTGGGCTGACGCCAAGCTCAGCTTCATCCACCGTGCCCAAC 739
QY 721 CTGGAGCGCTCGTTCCGGCAGAGGTGTGGAGGCGAGCTGACCGCACCCCTTCGCCCTG 780
Db 740 CTGGAGCGCTCGTTCCGGCAGAGGTGTGGAGGCGAGCTGACCGCACCCCTTCGCCCTG 799
QY 781 ACGCTCTCGGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCCATCCATCCCTGC 840
Db 800 ACGCTCTCGGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCCATCCATCCCTGC 859
QY 841 AACAGCGCACTGGGGGGAAGAGAGATCTGTAGTGCCCTTACTCAACCCATGGAC 900
Db 860 AACAGCGCACTGGGGGGAAGAGAGATCTGTAGTGCCCTTACTCAACCCATGGAC 919
QY 901 ATCCAGGTGCTCAGCAGGAGCGGAGCTTTCTTCCACACTCGCTGTGAGGAGACAAT 960
Db 920 ATCCAGGTGCTCAGCAGGAGCGGAGCTTTCTTCCACACTCGCTGTGAGGAGACAAT 979
QY 961 GCGGCTGCTCCCACTGTGCTGTCTGCCCAAGCGAGCCTTTCTACACATGCGCCTGC 1020
Db 980 GCGGCTGCTCCCACTGTGCTGTCTGCCCAAGCGAGCCTTTCTACATGCGCCTGC 1039
QY 1021 CCCACGGGTGTGAGCTGACGAGCAACCGGACGACGTGTAAAGGACAGGAGCGGAGAGGTG 1080
Db 1040 CCCACGGGTGTGAGCTGACGAGCAACCGGACGACGTGTAAAGGACAGGAGCGGAGAGGTG 1099
QY 1081 CTGCTGCTGGCCCGCGAGCGACCTACGAGGATCTCGCTGGACACGCCGAGCTTCACC 1140
Db 1100 CTGCTGCTGGCCCGCGAGCGACCTACGAGGATCTCGCTGGACACGCCGAGCTTCACC 1159
QY 1141 GACATCTGCTGCAAGTGGACGACATCCGGCAGCGCATTTGCCATCGACTAGACCCGCTA 1200
Db 1160 GACATCTGCTGCAAGTGGACGACATCCGGCAGCGCATTTGCCATCGACTAGACCCGCTA 1219
QY 1201 GAGGCTATGTCTACTGACAGATGACGAGTGGGGCCATCCGACAGCGGTACTCTGGAC 1260
Db 1220 GAGGCTATGTCTACTGACAGATGACGAGTGGGGCCATCCGACAGCGGTACTCTGGAC 1279
QY 1261 GGGTCTGGGGCGCAGACCTGGTCAACACGAGATCAACGACCCCGATGGCATCGCGTC 1320
Db 1280 GGGTCTGGGGCGCAGACCTGGTCAACACGAGATCAACGACCCCGATGGCATCGCGTC 1339
QY 1321 GACTGGGTGGCCGAAACCTTCTACTGGACCGACACGGGCAACCGCATCGAGGTGACG 1380
Db 1340 GACTGGGTGGCCGAAACCTTCTACTGGACCGACACGGGCAACCGCATCGAGGTGACG 1399
QY 1381 CGCTCAACGGGACCTCCGCGAAGATCTCTGTGTGCGAGGACCTTGGACGAGCGCCGAGCC 1440
Db 1400 CGCTCAACGGGACCTCCGCGAAGATCTCTGTGTGCGAGGACCTTGGACGAGCGCCGAGCC 1459
QY 1441 ATCGCACTGCACCCCGTGCATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAA 1500
Db 1460 ATCGCACTGCACCCCGTGCATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAA 1519
QY 1501 ATCGAGTGTGCCAACTTGGATGGGAGGAGCGGCGTGTGGTCAATGCCCTCCCTCGGG 1560
Db 1520 ATCGAGTGTGCCAACTTGGATGGGAGGAGCGGCGTGTGGTCAATGCCCTCCCTCGGG 1579
QY 1561 TGGCCCAACGGGCTGGCCCTGACCTGACGAGGAGGAGGCTTACTTGGGGAGAGCGCAAG 1620
Db 1580 TGGCCCAACGGGCTGGCCCTGACCTGACGAGGAGGAGGAGCTTACTTGGGGAGAGCGCAAG 1639
QY 1621 ACAGACAAGATCGAGGTGATCAATGTTGATGGGAGGAGGAGCGGACCTCTCGGAGGAC 1680
Db 1640 ACAGACAAGATCGAGGTGATCAATGTTGATGGGAGGAGGAGCGGACCTCTCGGAGGAC 1699
QY 1681 AAGCTCCCGCACATTTTCGGGTTCACGCTGCTGGGGGACTTTCATCTACTGGAAGTACTGG 1740
Db 1700 AAGCTCCCGCACATTTTCGGGTTCACGCTGCTGGGGGACTTTCATCTACTGGAAGTACTGG 1759
QY 1741 CAGCGCCGACGATCGAGCGGTGACAAAGTCAAGGCCAGCGGAGCTCATCATTTGAC 1800
Db 1760 CAGCGCCGACGATCGAGCGGTGACAAAGTCAAGGCCAGCGGAGCTCATCATTTGAC 1819

QY 1801 CAGCTGCCCGACCTGATGGGCTCAAGCTCTGAATGTGGCCAAAGTCTGTGGAGACCAAC 1860
Db 1820 CAGTGCCTCCGACCTGATGGGCTCAAGCTCTGAATGTGGCCAAAGTCTGTGGAGACCAAC 1879
QY 1861 CCGTGTGCGGACAGGAAACGGGGGTGACACCACTGTCTTTCTTCAACCCACCAACCAAC 1920
Db 1880 CCGTGTGCGGACAGGAAACGGGGGTGACACCACTGTCTTTCTTCAACCCACCAACCAAC 1939
QY 1921 CCGTGTGCTGCCCATCGGCTCGAGTGTGTGAGTGAATGAAGACTGTGATCTGTGCT 1980
Db 1940 CCGTGTGCTGCCCATCGGCTCGAGTGTGTGAGTGAATGAAGACTGTGATCTGTGCT 1999
QY 1981 GAGGCTCTTGTGCTTCAACAGCAGAGCGGCGCATCCACAGATCTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTCTTGTGCTTCAACAGCAGAGCGGCGCATCCACAGATCTCCCTCGAGACCAAT 2059
QY 2041 AACAAACGAGTGGGCTATCCCGCTCAACGGGCGTCAAGGAGGCTCAGCGCTCGACTTTGAT 2100
Db 2060 AACAAACGAGTGGGCTATCCCGCTCAACGGGCGTCAAGGAGGCTCAGCGCTCGACTTTGAT 2119
QY 2101 GTGTCCAAACACACATCTACTGACAGACGTCAGCTGAGCTGAAGACCATCAGCGCGCTTC 2160
Db 2120 GTGTCCAAACACACATCTACTGACAGACGTCAGCTGAGCTGAAGACCATCAGCGCGCTTC 2179
QY 2161 ATGAACGGGAGCTCGGTGGAGCAGTGTGTGAGTTTGGCCCTTGACTACCCGAGGCGATG 2220
Db 2180 ATGAACGGGAGCTCGGTGGAGCAGTGTGTGAGTTTGGCCCTTGACTACCCGAGGCGATG 2239
QY 2221 GCCGTGTGACTGGATGGGCAAGAACCTCTACTCTGGGCGGACACTGGGACCAACAGAAATCGAA 2280
Db 2240 GCCGTGTGACTGGATGGGCAAGAACCTCTACTCTGGGCGGACACTGGGACCAACAGAAATCGAA 2299
QY 2281 GTGGCGGCTGGACGGGCAAGTTCCGGCAAGTCTCTGTGTGAGGAGACTTGGACAAACCG 2340
Db 2300 GTGGCGGCTGGACGGGCAAGTTCCGGCAAGTCTCTGTGTGAGGAGACTTGGACAAACCG 2359
QY 2341 AGGTCTGCTGGCCCTGGATCCCAACAGGCTTACATCTACTGACCGAGTGGGGGCGAAG 2400
Db 2360 AGGTCTGCTGGCCCTGGATCCCAACAGGCTTACATCTACTGACCGAGTGGGGGCGAAG 2419
QY 2401 CCGAGGATCTGTCGGGCTTCAATGACCGGACCAACTGATGACGCTGGTGAGCAAGGTG 2460
Db 2420 CCGAGGATCTGTCGGGCTTCAATGACCGGACCAACTGATGACGCTGGTGAGCAAGGTG 2479
QY 2461 GGGCGGGCAACGACCTCAACATTTGACTACGCTGACCGCGCTTACTTGGACCGACCTG 2520
Db 2480 GGGCGGGCAACGACCTCAACATTTGACTACGCTGACCGCGCTTACTTGGACCGACCTG 2539
QY 2521 GACACCAACATGATCGAGTCTGTCACCATGCTGGGTGAGGAGCGGCTGTGATTTGCCGAC 2580
Db 2540 GACACCAACATGATCGAGTCTGTCACCATGCTGGGTGAGGAGCGGCTGTGATTTGCCGAC 2599
QY 2581 GATCTCCCGCAACCCGTTTGGTCTGACCGAGTACGAGCTTATATCTACTGACAGACTGG 2640
Db 2600 GATCTCCCGCAACCCGTTTGGTCTGACCGAGTACGAGCTTATATCTACTGACAGACTGG 2659
QY 2641 AATCTGACAGACTTGGCGGCGGACCAAGACTAGCGGCGGAGACCGACCTCATTCAG 2700
Db 2660 AATCTGACAGACTTGGCGGCGGACCAAGACTAGCGGCGGAGACCGACCTCATTCAG 2719
QY 2701 GGGCACCTGGAATTTCTGTGATGGACATCTGTGTGTTTCACTCTCCCGCAGGATGGCCTC 2760
Db 2720 GGGCACCTGGAATTTCTGTGATGGACATCTGTGTGTTTCACTCTCCCGCAGGATGGCCTC 2779
QY 2761 AATGACTGTATGCAACAAACGGGAGTGTGGGAGTGTGCTTGGCATCCCGGCGGCG 2820
Db 2780 AATGACTGTATGCAACAAACGGGAGTGTGGGAGTGTGCTTGGCATCCCGGCGGCG 2839
QY 2821 CACCGCTGGGCTGGCTCAGCTACACTGACCTGACCCCGCAGCGCACTGACGAGCCG 2880
Db 2840 CACCGCTGGGCTGGCTCAGCTACACTGACCTGACCCCGCAGCGCACTGACGAGCCG 2899

Db 5060 GGCAGGGCTGGAGAACTTTGTA 5082

RESULT 9

ADR16919

ADR16919 standard; DNA; 5120 BP.

AC ADR16919;

DT 04-NOV-2004 (first entry)

DE Human high bone mass gene, wild type allele Zmax1 DNA.

Human; high bone mass; Zmax1; ds; gene; HBM; osteoporosis;
KW
chromosome 11q13.3; osteopathic; LDL receptor; bone development;
KW
metabolic bone disease; SNP; single nucleotide polymorphism.
KW

44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562

AA	Key	Location/Qualifiers
FH	CDS	71. .4918
FT		

```

FT      tag="a"
FT      /product="Zmax1 protein"
FT      /transl_except="pos:530..532,aa:Lys)
FT      /note="Applies to SEQ ID 3 only"
FT      /transl_except="pos:986..988,aa:Trp)
FT      /note="Applies to SEQ ID 3 only"
FT      /transl_except="pos:1055..1057,aa:Met)
FT      /note="Applies to SEQ ID 3 only"
FT      /transl_except="pos:2162..2164,aa:Asn)
FT      /note="Applies to SEQ ID 3 only"
FT      replace(582,T
FT      allele

```

PN US6780609-B1.

24-AUG-2004:

05-APR-2000; 2000US-00543771.

PR 13-JAN-1998; 98US-0071449P.

PR 23-OCT-1998; 98US-0105511P.

PR 13-JAN-1999; 99US-00229319.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Carulli JP, Little RD, Recker RR, Johnson ML;

DR WPI; 2004-623529/60.

DR P-PSDB; ADR16921, ADR17560.

PT New high bone mass gene of chromosome 1.1Q13.3, encoding protein useful for treating, diagnosing, preventing, or screening for normal and PT abnormal conditions of bone, including metabolic bone diseases, e.g. PT osteoporosis.

PS Disclosure: SEQ ID NO 1: 284pp: English.

The invention relates to an isolated amino acid protein sequence selected from an amino acid sequence appearing as ADR16922 or an amino acid sequence comprising or consisting of the extracellular domain of ADR16922(amino acids 23-1395). ADR16922 is encoded by the HBM (high bone mass) allele of the human *Xmai1* gene and has sequence similarity to LDL receptors. Also disclosed are nucleic acids, proteins, cloning vectors, expression vectors, transformed hosts, methods of developing pharmaceutical compositions, methods of identifying molecules involved in bone development, and methods of diagnosing and treating diseases involved in bone development. Specifically disclosed is the *Xmai1* gene and the high bone mass (HBM) allele on chromosome 11q13.3 encoding ADR16922. The protein is useful for treating, diagnosing, preventing, or screening for normal and abnormal conditions of bone, including metabolic bone diseases, e.g. osteoporosis. The present sequence is the wild type

Qy 961 GCGGCTGCTCCACCTGTGCTGCTGCCCAAGCGAGCCTTTCTACACATGCGCCTGC 1020
Db 980 GCGGCTGCTCCACCTGTGCTGCTGCCCAAGCGAGCCTTTCTACACATGCGCCTGC 1039
Qy 1021 CCCAGGCTGTGAGCTGCAGGACAAACGGCAGGAGCTGTAAAGGACGAGGAGGAGG 1080
Db 1040 CCCAGGCTGTGAGCTGCAGGACAAACGGCAGGAGCTGTAAAGGACGAGGAGGAGG 1099
Qy 1081 CTGCTGCTGGCCCGCGGACGAGCCTACGAGGAGATCTCGCTGAGACACGCCGAGCTTCACC 1140
Db 1100 CTGCTGCTGGCCCGCGGACGAGCCTACGAGGAGATCTCGCTGAGACACGCCGAGCTTCACC 1159
Qy 1141 GACATGCTGCTGAGGCTGAGACGACATCCGGCAGCGCAATTGGCATCGACTACGAGCCGCTA 1200
Db 1160 GACATGCTGCTGAGGCTGAGACGACATCCGGCAGCGCAATTGGCATCGACTACGAGCCGCTA 1219
Qy 1201 GAGGCTATGTCTACTGAGACAGATGACGAGGTGGGGCCATCCGAGGCGGTACTCTGAGC 1260
Db 1220 GAGGCTATGTCTACTGAGACAGATGACGAGGTGGGGCCATCCGAGGCGGTACTCTGAGC 1279
Qy 1261 GGGTCTGGGGCGAGACGCTGCTCAACACCGAGATCAACGACCCCGATGGCATCGCGGTC 1320
Db 1280 GGGTCTGGGGCGAGACGCTGCTCAACACCGAGATCAACGACCCCGATGGCATCGCGGTC 1339
Qy 1321 GACTGGGTGGCCCGAAACCTCTACTGAGCCGACACGGGACCGCATCGAGGTGACG 1380
Db 1340 GACTGGGTGGCCCGAAACCTCTACTGAGCCGACACGGGACCGCATCGAGGTGACG 1399
Qy 1381 GCGCTCAACGGGACCTCCGCAAGATCTGTGTGCGAGGACCTTGAGAGCCCGCGAGCC 1440
Db 1400 GCGCTCAACGGGACCTCCGCAAGATCTGTGTGCGAGGACCTTGAGAGCCCGCGAGCC 1459
Qy 1441 ATCCGACTGACCCCGTGATGGGCTCATGTACTGAGCAGACTGGGAGAGAACCTTAAA 1500
Db 1460 ATCCGACTGACCCCGTGATGGGCTCATGTACTGAGCAGACTGGGAGAGAACCTTAAA 1519
Qy 1501 ATCGAGTGTGCAACTTTGGATGGGCGAGGAGGGGTGTGCTGGTCAATGCTCCCTCGGG 1560
Db 1520 ATCGAGTGTGCAACTTTGGATGGGCGAGGAGGGGTGTGCTGGTCAATGCTCCCTCGGG 1579
Qy 1561 TGGCCCAACGGCCTGGCCCTGGACCTTGACGAGGAGGAGAGCTCTACTGGGAGAGACCGAAG 1620
Db 1580 TGGCCCAACGGCCTGGCCCTGGACCTTGACGAGGAGGAGAGCTCTACTGGGAGAGACCGAAG 1639
Qy 1621 ACAGCAAGATCGAGGTGATCAATGTTGATGCGAGCGAGGCGGACCCCTCTCGAGGAC 1680
Db 1640 ACAGCAAGATCGAGGTGATCAATGTTGATGCGAGCGAGGCGGACCCCTCTCGAGGAC 1699
Qy 1681 AAGCTCCCGCACATTTTCGGGTTTCACTGCTGGGGGACTTTCATCTACTGGACTGACTGG 1740
Db 1700 AAGCTCCCGCACATTTTCGGGTTTCACTGCTGGGGGACTTTCATCTACTGGACTGACTGG 1759
Qy 1741 CAGCGCGCAGCATCGAGCGGGTGCACAGGTCAAGGCACGCGGAGAGTCAATCATTTGAC 1800
Db 1760 CAGCGCGCAGCATCGAGCGGGTGCACAGGTCAAGGCACGCGGAGAGTCAATCATTTGAC 1819
Qy 1801 CAGCTGCGCAGCTGATGGGCTCAAGCTGCTGATGCGCAAGGTGCTCGGAACCAAC 1860
Db 1820 CAGCTGCGCAGCTGATGGGCTCAAGCTGCTGATGCGCAAGGTGCTCGGAACCAAC 1879
Qy 1861 CCGTGTGCGGACAGGAAACGGGGGTGACGCCACTGTGCTTCTTTCACACCCCAAGCAACC 1920
Db 1880 CCGTGTGCGGACAGGAAACGGGGGTGACGCCACTGTGCTTCTTTCACACCCCAAGCAACC 1939
Qy 1921 CCGTGTGCTGCCCATCGGCTCGAGCTGTGATGATGACATGAAGACCTTGATCGTGCTT 1980
Db 1940 CCGTGTGCTGCCCATCGGCTCGAGCTGTGATGATGACATGAAGACCTTGATCGTGCTT 1999
Qy 1981 GAGGCTTCTTGTGCTTTCACAGCAGAGCGCGCATCCACAGGATCTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTTCTTGTGCTTTCACAGCAGAGCGCGCATCCACAGGATCTCCCTCGAGACCAAT 2059
Qy 2041 AACCAACGACGTGGCCATCCCGCTCAACGGGCGCTCAAGGAGGCGCTCAGGCCCTGGACTTTGAT 2100

Db 2060 AACCAACGACGTGGCCATCCCGCTCAACGGGCGCTCAAGGAGGCGCTCAGGCCCTGGACTTTGAT 2119
Qy 2101 GTGTCCAAACAAACACATCTACTGACAGACGCTCAGCCTGAAAGACCATCAGCGCGGCTTC 2160
Db 2120 GTGTCCAAACAAACACATCTACTGACAGACGCTCAGCCTGAAAGACCATCAGCGCGGCTTC 2179
Qy 2161 ATGAACGGGAGCTCGGTGGAGCAGTGTGTGAGATTTGGCCCTTGACTACCCCGAGGGCATG 2220
Db 2180 ATGAACGGGAGCTCGGTGGAGCAGTGTGTGAGATTTGGCCCTTGACTACCCCGAGGGCATG 2239
Qy 2221 GCCGTTGACTGGATGGGCAAGAACCTCTACTTGGGCGGACACTTGGGACCAACAGATCGAA 2280
Db 2240 GCCGTTGACTGGATGGGCAAGAACCTCTACTTGGGCGGACACTTGGGACCAACAGATCGAA 2299
Qy 2281 GTGGCGGGCTGGAGCGGGCAGTTCGGCAAGTCTCTGTGTGAGGGACTTGGACAACCCG 2340
Db 2300 GTGGCGGGCTGGAGCGGGCAGTTCGGCAAGTCTCTGTGTGAGGGACTTGGACAACCCG 2359
Qy 2341 AGTTCGCTGGGCCCTGGATCCCAACAAAGGCTTACTTCTGAGCCGAGTGGGGCGGCAAG 2400
Db 2360 AGTTCGCTGGGCCCTGGATCCCAACAAAGGCTTACTTCTGAGCCGAGTGGGGCGGCAAG 2419
Qy 2401 CCGAGGATCGTGGCGGCTTTCATGGACGGGACCAACTGTCATGACGCTGGTGGACAAGGTG 2460
Db 2420 CCGAGGATCGTGGCGGCTTTCATGGACGGGACCAACTGTCATGACGCTGGTGGACAAGGTG 2479
Qy 2461 GCGCGGGCAACGACCTTCAACATTTGACTACGCTGACACGCGCTTCTACTGAGCAGCCTG 2520
Db 2480 GCGCGGGCAACGACCTTCAACATTTGACTACGCTGACACGCGCTTCTACTGAGCAGCCTG 2539
Qy 2521 GACACCAACATGATCGAGTGTGTCACCAATGCTGCTGAGGAGCGGCTGTGATTTGCCGAC 2580
Db 2540 GACACCAACATGATCGAGTGTGTCACCAATGCTGCTGAGGAGCGGCTGTGATTTGCCGAC 2599
Qy 2581 GATCTCCCGCACCCGTTTCGCTGTCGACGAGTACAGCGATTAATATCTACTTGGACAGACTGG 2640
Db 2600 GATCTCCCGCACCCGTTTCGCTGTCGACGAGTACAGCGATTAATATCTACTTGGACAGACTGG 2659
Qy 2641 AATCTGACAGCATTGAGCGGGCCGACAGACTAGCGGCGGAAACCGACCTCATCCAG 2700
Db 2660 AATCTGACAGCATTGAGCGGGCCGACAGACTAGCGGCGGAAACCGACCTCATCCAG 2719
Qy 2701 GGCACCTGACTTCTGTGATGGACATCTGTGTGTTTCACTCTCCCGCAGAGTGGCCTC 2760
Db 2720 GGCACCTGACTTCTGTGATGGACATCTGTGTGTTTCACTCTCCCGCAGAGTGGCCTC 2779
Qy 2761 AATGACTGTATGCAAAACAGGGCAGTGTGGGAGCTGTGCTTGGCATCCCGGGCGGC 2820
Db 2780 AATGACTGTATGCAAAACAGGGCAGTGTGGGAGCTGTGCTTGGCATCCCGGGCGGC 2839
Qy 2821 CACGCTGCGGCTGGCCTCACACTACACCTGGAACCCAGCAGCGGCACTGACGCCG 2880
Db 2840 CACGCTGCGGCTGGCCTCACACTACACCTGGAACCCAGCAGCGGCACTGACGCCG 2899
Qy 2881 CCCACACCTTCTGTGCTTTCAGCCAGAAATCTCCCATCAGTGTGATGATCCCGAGCAGC 2940
Db 2900 CCCACACCTTCTGTGCTTTCAGCCAGAAATCTCCCATCAGTGTGATGATCCCGAGCAGC 2959
Qy 2941 CAGCACAGCCGGAATCTCATCTCTGCCCTTGATGGAGTGAAGAACGTCGATGAGGAGGAG 3000
Db 2960 CAGCACAGCCGGAATCTCATCTCTGCCCTTGATGGAGTGAAGAACGTCGATGAGGAGGAG 3019
Qy 3001 TATGACCCACTGGACAAGTTTCATCTACTTGGTGGATGGGGCCGAGAACATCAAGGAGCC 3060
Db 3020 TATGACCCACTGGACAAGTTTCATCTACTTGGTGGATGGGGCCGAGAACATCAAGGAGCC 3079
Qy 3061 AAGGACGAGCGGACCCGAGCCCTTTGTTTGTGACCTTCTGTAGCCAAAGGCGCAAAACCCAGAC 3120
Db 3080 AAGGACGAGCGGACCCGAGCCCTTTGTTTGTGACCTTCTGTAGCCAAAGGCGCAAAACCCAGAC 3139
Qy 3121 AGGAGGCCCAACGACCTCAGCATCGACATCTTACAGCCGGAACATCTGTTTCTTGGACGTGCGAG 3180

Qy	1381	CGCCTCAACGGGCACTCTCCGCAAGATCTGTGTGCGAGGACCTGGAGAGCCCGCGAGCC	1440
Db	1429	CGCCTCAACGGCACCTCCGCAAGATCTGTGTGCGAGGACCTGGAGAGCCCGCGAGCC	1488
Qy	1441	ATCGCACTGCAACCCGCTGATCGGCTCATGTACTGAGCAGACTGGGAGAGAACCTTAAA	1500
Db	1489	ATCGCACTGCAACCCGCTGATCGGCTCATGTACTGAGCAGACTGGGAGAGAACCTTAAA	1548
Qy	1501	ATCGAGTGTGCAACTTTGGATGGGAGGAGCGGCTGTGTGCTCAATGCTCCTCGGG	1560
Db	1549	ATCGAGTGTGCAACTTTGGATGGGAGGAGCGGCTGTGTGCTCAATGCTCCTCGGG	1608
Qy	1561	TGGCCCAACGCTGCTGCTGCACTGCAAGGGGAGCTTACTGGGAGAGCGCAAG	1620
Db	1609	TGGCCCAACGCTGCTGCTGCACTGCAAGGGGAGCTTACTGGGAGAGCGCAAG	1668
Qy	1621	ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAAGGCGGACCTCTGAGGAC	1680
Db	1669	ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAAGGCGGACCTCTGAGGAC	1728
Qy	1681	AAGCTCCGCACTTTTCGGGTTCACTGCTGCGGGGACTTCACTACTGGAAGTGTG	1740
Db	1729	AAGCTCCGCACTTTTCGGGTTCACTGCTGCGGGGACTTCACTACTGGAAGTGTG	1788
Qy	1741	CAGCGCCGACATCGAGCGGGTGCAAAAGTCAAGGCCAGCGGAGCTCATATTGAC	1800
Db	1789	CAGCGCCGACATCGAGCGGGTGCAAAAGTCAAGGCCAGCGGAGCTCATATTGAC	1848
Qy	1801	CAGCTGCCGCACTGATGGGCTCAAGCTGTGAATGGGCCAAGGTGCTCGGAACCAAC	1860
Db	1849	CAGCTGCCGCACTGATGGGCTCAAGCTGTGAATGGGCCAAGGTGCTCGGAACCAAC	1908
Qy	1861	CCGTGTGCGGACAGAAACGGGGGTGACGCACTGTGCTTTTCAACCCCAAGCAAC	1920
Db	1909	CCGTGTGCGGACAGAAACGGGGGTGACGCACTGTGCTTTTCAACCCCAAGCAAC	1968
Qy	1921	CGGTGTGCTGCCCATCGGCTGAGCTGTGATGATCATGAGACCTGATGCTGCT	1980
Db	1969	CGGTGTGCTGCCCATCGGCTGAGCTGTGATGATCATGAGACCTGATGCTGCT	2028
Qy	1981	GAGGCTTTCTTGGTCTTCAACAGCAGAGCCGCGCATCCACAGGATCTCCCTCGAGACCAAT	2040
Db	2029	GAGGCTTTCTTGGTCTTCAACAGCAGAGCCGCGCATCCACAGGATCTCCCTCGAGACCAAT	2088
Qy	2041	AACAAACGAGTGGCCATCCCGCTCAAGGCGTCAAGGAGGCTCAGCCCTGGAATTGAT	2100
Db	2089	AACAAACGAGTGGCCATCCCGCTCAAGGCGTCAAGGAGGCTCAGCCCTGGAATTGAT	2148
Qy	2101	GTGTCCAAACCAACATCTACTGACAGAGCTCAGCCTGAGAGACCATCAGCGCGCTTC	2160
Db	2149	GTGTCCAAACCAACATCTACTGAGACAGCTCAGCCTGAGAGACCATCAGCGCGCTTC	2208
Qy	2161	ATGAACGGGAGCTCGGTGGAGCACTGTGTGAGTTTGGCTTGAATACCCCGAGGGCATG	2220
Db	2209	ATGAACGGGAGCTCGGTGGAGCACTGTGTGAGTTTGGCTTGAATACCCCGAGGGCATG	2268
Qy	2221	GGCGTTGATGATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGATCGAA	2280
Db	2269	GGCGTTGATGATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGATCGAA	2328
Qy	2281	GTGGCGGCTGGACGGGCAAGTTCCGGCAAGTCTCTGTGTGAGGGAATTGGACAAACCG	2340
Db	2329	GTGGCGGCTGGACGGGCAAGTTCCGGCAAGTCTCTGTGTGAGGGAATTGGACAAACCG	2388
Qy	2341	AGGTGCTGGCCCTGGATCCCAACGAGGCTCATCTACTGGAACGAGTGGGGCGGCAAG	2400
Db	2389	AGGTGCTGGCCCTGGATCCCAACGAGGCTCATCTACTGGAACGAGTGGGGCGGCAAG	2448
Qy	2401	CCGAGGATGCTCGGGCTTCAATGAGCGGACCAACTGATGACCTGGTGACAGGTTG	2460
Db	2449	CCGAGGATGCTCGGGCTTCAATGAGCGGACCAACTGATGACCTGGTGACAGGTTG	2508
Qy	2461	GGCGGGGCAACGACCTCAACCATGACTACGTGACCGAGCGCTCTACTGGACCGACCTG	2520

Db	2509	GGCGGGCCAAAGACCTTCAACATTGACTACGTGACCGAGGCCCTCTACTGGACCGACTG	2568
Qy	2521	GACACCAACATGATCGAGTCTGTCAACATCTGTGGTTCAGAGCGGGTCTGTGATTCGCGAC	2580
Db	2569	GACACCAACATGATCGAGTCTGTCAACATCTGTGGTTCAGAGCGGGTCTGTGATTCGCGAC	2628
Qy	2581	GATCTCCGACACCGTTCTGGTCTGAGCCAGTACAGCGATTATATCTACTTGGACAGACTGG	2640
Db	2629	GATCTCCGACACCGTTCTGGTCTGAGCCAGTACAGCGATTATATCTACTTGGACAGACTGG	2688
Qy	2641	AATCTGCACAGCATTTGAGCGGGCCGACAAAGACTAGCCGGCCGGAACCGCACCTCATCCAG	2700
Db	2689	AATCTGCACAGCATTTGAGCGGGCCGACAAAGACTAGCCGGCCGGAACCGCACCTCATCCAG	2748
Qy	2701	GGCCACTGGAATTCGTGTGATGGACATCTCTGTGTTCATCTCTCCCGCAGGATGGGCTC	2760
Db	2749	GGCCACTGGAATTCGTGTGATGGACATCTCTGTGTTCATCTCTCCCGCAGGATGGGCTC	2808
Qy	2761	AATGACTGTATGCACAAAGGGGAGTGTGGGAGCTGTGCTTGGCATCCCGGGGGC	2820
Db	2809	AATGACTGTATGCACAAAGGGGAGTGTGGGAGCTGTGCTTGGCATCCCGGGGGC	2868
Qy	2821	CACCGCTGGGCTGGGCTCACACTACACCTTGACCCCGACAGCCGCAACTGCGAGCCG	2880
Db	2869	CACCGCTGGGCTGGGCTCACACTACACCTTGACCCCGACAGCCGCAACTGCGAGCCG	2928
Qy	2881	CCCAACCACTTTGCTGTTCAGCCAGAAATCTGCCATCTAGTCGGATGATCCCGGACGAC	2940
Db	2929	CCCAACCACTTTGCTGTTCAGCCAGAAATCTGCCATCTAGTCGGATGATCCCGGACGAC	2988
Qy	2941	CAGCACAGCCGGATCTCATCTGCCCCCTGATGGACTGAGGAACTGAAAGCCATCGAC	3000
Db	2989	CAGCACAGCCGGATCTCATCTGCCCCCTGATGGACTGAGGAACTGAAAGCCATCGAC	3048
Qy	3001	TATGACCCACTGGACAAAGTTTCTACTTGGTGTGATGGGCGCCAGAAACATCAAGCGAGCC	3060
Db	3049	TATGACCCACTGGACAAAGTTTCTACTTGGTGTGATGGGCGCCAGAAACATCAAGCGAGCC	3108
Qy	3061	AAGGACGAGCGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAAAGCCAAACCCAGAC	3120
Db	3109	AAGGACGAGCGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAAAGCCAAACCCAGAC	3168
Qy	3121	AGGACGCGCCACGACCTCAGCATCGACATCTTACAGCCGACACTGTCTTGACGCTGCGAG	3180
Db	3169	AGGACGCGCCACGACCTCAGCATCGACATCTTACAGCCGACACTGTCTTGACGCTGCGAG	3228
Qy	3181	GCCACCAATACCATCAACGCTCCACAGGCTGAGCGGGGAAGCCATGGGGGTGGTGTGCTG	3240
Db	3229	GCCACCAATACCATCAACGCTCCACAGGCTGAGCGGGGAAGCCATGGGGGTGGTGTGCTG	3288
Qy	3241	GGGACCGCGACAAAGCCAGGGCCATCGTCTCAACGGGAGCGAGGCTACTGTACTTC	3300
Db	3289	GGGACCGCGACAAAGCCAGGGCCATCGTCTCAACGGGAGCGAGGCTACTGTACTTC	3348
Qy	3301	ACCAACATGACGAGCGGGCAGCCAAAGATCGAACCGCAGCCCTGAGCGCACCGAGCGC	3360
Db	3349	ACCAACATGACGAGCGGGCAGCCAAAGATCGAACCGCAGCCCTGAGCGCACCGAGCGC	3408
Qy	3361	GAGGTCTCTTTCACCAACCGGCTCATCGCCCTGTGGCCCTGGTGGTGGACAAACACTG	3420
Db	3409	GAGGTCTCTTTCACCAACCGGCTCATCGCCCTGTGGCCCTGGTGGTGGACAAACACTG	3468
Qy	3421	GGCAAGCTGTTCTGGGTGGACCGGACCTGAAAGCGCATGAGAGCTGTGACTGTGTCAGG	3480
Db	3469	GGCAAGCTGTTCTGGGTGGACCGGACCTGAAAGCGCATGAGAGCTGTGACTGTGTCAGG	3528
Qy	3481	GCCAAACCGCTGACCTCTGGAGGACCGCAACATCTGTGACGCTCTGGGGCTGACCACTCTT	3540
Db	3529	GCCAAACCGCTGACCTCTGGAGGACCGCAACATCTGTGACGCTCTGGGGCTGACCACTCTT	3588
Qy	3541	GGCAAGCACTCTTACTGTGATCGACCGCCAGCAGCAGATGATCGAGCGTGTGGAGAGACC	3600

Db 3589 GCGAAGCATCTTCTACTGGATCGACCGCCAGCAGCAGATGATCGAGCGTGTGGAGAAGACC 3648
Qy 3601 ACCGGGGAACAGCGACTCGCATCCAGGCGCGTGTGCGCCACCTCACTGGCATCCATGCA 3660
Db 3649 ACCGGGGAACAGCGACTCGCATCCAGGCGCGTGTGCGCCACCTCACTGGCATCCATGCA 3708
Qy 3661 GTGGAGGAAGTCAAGCTCGAGGAGTTCTCAGCCCAACCATGTGCGCCGTGACAATGGTGGC 3720
Db 3709 GTGGAGGAAGTCAAGCTCGAGGAGTTCTCAGCCCAACCATGTGCGCCGTGACAATGGTGGC 3768
Qy 3721 TGCTCCCAATCTGTATTTGCCAAGGGTGATGGGACACACAGGTGCTCATGCGCATGCCAC 3780
Db 3769 TGCTCCCAATCTGTATTTGCCAAGGGTGATGGGACACACAGGTGCTCATGCGCATGCCAC 3828
Qy 3781 CTGCTGCTCTCGCAAACTGTGACCTGTGAGAGCGCCACCTGTCTCCCGGACACAG 3840
Db 3829 CTGCTGCTCTCGCAAACTGTGACCTGTGAGAGCGCCACCTGTCTCCCGGACACAG 3888
Qy 3841 TTTGCATGTGCCACAGGGAGATCGACTGTATCTCCCGGGGCTGGCGCTGTGACGGCTTT 3900
Db 3889 TTTGCATGTGCCACAGGGAGATCGACTGTATCTCCCGGGGCTGGCGCTGTGACGGCTTT 3948
Qy 3901 CCCGAGTGGATGACACAGAGCGACGAGAGGGTGCCTCCCGTGTGCTCCCGCCCGCAGTTC 3960
Db 3949 CCCGAGTGGATGACACAGAGCGACGAGAGGGTGCCTCCCGTGTGCTCCCGCCCGCAGTTC 4008
Qy 3961 CCTGCGCGCGGGGTCACTGTGTGACCTGTGCGCTGCGCTGCGAGCGGCGAGCGACTGT 4020
Db 4009 CCTGCGCGCGGGGTCACTGTGTGACCTGTGCGCTGCGCTGCGAGCGGCGAGCGACTGT 4068
Qy 4021 CAGGACCGCTCAGAGCGGAGTGTGACCGCATCTGCTGCCCAACAGTTCGGGTGT 4080
Db 4069 CAGGACCGCTCAGAGCGGAGTGTGACCGCATCTGCTGCCCAACAGTTCGGGTGT 4128
Qy 4081 CGGAGCGGCGAGTGTCTCATCAAACAGAGTGGCATCTCTTCCCGCAGTGTATCGAC 4140
Db 4129 CGGAGCGGCGAGTGTCTCATCAAACAGAGTGGCATCTCTTCCCGCAGTGTATCGAC 4188
Qy 4141 GGCTCCGAGAGTCACTGTGTGAATCACCAGCGCCCTCAGAGCAGCAGCGCGCCAC 4200
Db 4189 GGCTCCGAGAGTCACTGTGTGAATCACCAGCGCCCTCAGAGCAGCAGCGCGCCAC 4248
Qy 4201 AGCAGTGCATCGGCGCGCTCAITGGCATCATCTCTCTCTTCTGCTATGGGTGGTGTG 4260
Db 4249 AGCAGTGCATCGGCGCGCTCAITGGCATCATCTCTCTCTTCTGCTATGGGTGGTGTG 4308
Qy 4261 TATTTTGTGTCCAGCGGTGTGTGCGAGCGTATGCGGGGGCAACGCGGCGCTTCCCG 4320
Db 4309 TATTTTGTGTCCAGCGGTGTGTGCGAGCGTATGCGGGGGCAACGCGGCGCTTCCCG 4368
Qy 4321 CACGAGTATGTCAGCGGACCCCGCAGCTGCGCCCTCAATTTTCATAGCCCGGGCGGTTC 4380
Db 4369 CACGAGTATGTCAGCGGACCCCGCAGCTGCGCCCTCAATTTTCATAGCCCGGGCGGTTC 4428
Qy 4381 CAGCATGGCCCTTTCACAGGATCGCATCGGAAAGTCCATGATGAGTCCCGTGAGCGCTG 4440
Db 4429 CAGCATGGCCCTTTCACAGGATCGCATCGGAAAGTCCATGATGAGTCCCGTGAGCGCTG 4488
Qy 4441 ATGGGGGGCGGGGGGGTGCCTCTGTACACCGGAACACGTCACAGGGGCTCGTTC 4500
Db 4489 ATGGGGGGCGGGGGGGTGCCTCTGTACACCGGAACACGTCACAGGGGCTCGTTC 4548
Qy 4501 AGCAGTGTGTCAGACGAGCGCAGCTGTATCCCGCGATCTGACCCCGCGCCCTCC 4560
Db 4549 AGCAGTGTGTCAGACGAGCGCAGCTGTATCCCGCGATCTGACCCCGCGCCCTCC 4608
Qy 4561 CCGGCGACGAGCCCTCCCTGTACAAACATGACATGTTTCTACTCTTCAAACTTCGCGCC 4620
Db 4609 CCGGCGACGAGCCCTCCCTGTACAAACATGACATGTTTCTACTCTTCAAACTTCGCGCC 4668
Qy 4621 ACTGCGAGACGTCAGAGCCCTTACATCATTCAGGAATGGCGCCCGCAGACGCCCTTCG 4680
Db 4669 ACTGCGAGACGTCAGAGCCCTTACATCATTCAGGAATGGCGCCCGCAGACGCCCTTCG 4728

Qy 4681 AGCACCGACGTTGTGACAGCGACTACAGCGCCAGCGCTGGAAGGCCAGCAAGTACTAC 4740
Db 4729 AGCACCGACGTTGTGACAGCGACTACAGCGCCAGCGCTGGAAGGCCAGCAAGTACTAC 4788
Qy 4741 CTGGATTTGAATCTGGACTCAGACCCCTATCCACCCCAACCCCAACCCCAACCCCAAGTATC 4800
Db 4789 CTGGATTTGAATCTGGACTCAGACCCCTATCCACCCCAACCCCAACCCCAACCCCAAGTATC 4848
Qy 4801 CTGTGCGGAGGACAGCTGCGCGCCCTGCGCGCCCAACCCCAACCCCAACCCCAAGTATC 4860
Db 4849 CTGTGCGGAGGACAGCTGCGCGCCCTGCGCGCCCAACCCCAACCCCAACCCCAAGTATC 4908
Qy 4861 TTCCGCGCCCTCTCGCTGCGCCCTGCGCGCCCAACCCCAACCCCAACCCCAAGTATC 4920
Db 4909 TTCCGCGCCCTCTCGCTGCGCCCTGCGCGCCCAACCCCAACCCCAACCCCAAGTATC 4968
Qy 4921 TTCTCTGTGCGCCCTGTAAATAGTTTAAATATGAACAAAGAAATAATATATATATATATAT 4980
Db 4969 TTCTCTGTGCGCCCTGTAAATAGTTTAAATATGAACAAAGAAATAATATATATATATATAT 5028
Qy 4981 TTAAATAATAATAATATGCGATTTTAAATAATGAACAAAGAAATAATATATATATATATAT 5040
Db 5029 TTAAATAATAATAATATGCGATTTTAAATAATGAACAAAGAAATAATATATATATATATAT 5088
Qy 5041 GGCAGGGCTGGGAGAACTTTCTGA 5063
Db 5089 GGCAGGGCTGGGAGAACTTTCTGA 5111

RESULT 11

ADB98797
ID ADB98797 standard; DNA; 5278 BP.

XX ADB98797;

XX 04-DEC-2003 (first entry)

XX Human ZMAXGI-3AS DNA sequence.

XX Osteopathic; Gene therapy; High Bone Mass; HBM; LRP5; Zmax1; LRP6;

XX bone mass modulation; osteoporosis; human; ds.

XX Homo sapiens.

XX WO200292000-A2.

XX 21-NOV-2002.

XX 13-MAY-2002; 2002WO-US014877.

XX 11-MAY-2001; 2001US-0290071P.

XX 17-MAY-2001; 2001US-0291311P.

XX 01-FEB-2002; 2002US-0353058P.

XX 04-MAR-2002; 2002US-0361293P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX (AMHP) WYETH.

XX Allen K, Anisowicz A, Graham JR, Morales A, Yaworsky PJ, Liu W;

XX WPI; 2003-129214/12.

XX New nucleic acid comprising a mutation in LRP5 or LRP6, useful for

XX diagnosing a HBM-like phenotype in a subject and for preparing a

XX composition for modulating bone mass and/or lipid levels in a subject

XX suffering from e.g. osteoporosis.

XX Disclosure; Fig 25; 629pp; English.

XX The present invention relates to High Bone Mass (HBM), LRP5 (Zmax1) and

XX LRP6 mutants, which results in a HBM-like phenotype when expressed in a

XX cell. The HBM-like phenotype results in bone mass modulation and/or lipid

QY 2284 CGCGGCTGGACGGGCACTTCGGGCACTCTCGTGTGGAGGACTTGGACAAACCCGAGG 2343
DB 2281 CGCGGCTGGACGGGCACTTCGGGCACTCTCGTGTGGAGGACTTGGACAAACCCGAGG 2340
QY 2344 TCGCTGGCCCTTGGATCCCAACAAAGGGCTACACTTACTTGAACCGAGTGGGGCGGCAAGCCG 2403
DB 2341 TCGCTGGCCCTTGGATCCCAACAAAGGGCTACACTTACTTGAACCGAGTGGGGCGGCAAGCCG 2400
QY 2404 AGGATCTGTGGCGGCTTCATGGACGGGACCAACTGCTATGACGCTGTGGTGGACAAAGTGGGC 2463
DB 2401 AGGATCTGTGGCGGCTTCATGGACGGGACCAACTGCTATGACGCTGTGGTGGACAAAGTGGGC 2460
QY 2464 CGGCGCAACGACCTCACCATTTGACTAGCTGACGCGGCTCTACTTGAACCGAGCTGGAC 2523
DB 2461 CGGCGCAACGACCTCACCATTTGACTAGCTGACGCGGCTCTACTTGAACCGAGCTGGAC 2520
QY 2524 ACCAATGATGATGATGCTCCAAATGCTGGGTGAGGAGCGGGTCTGATTTGCCGACGAT 2583
DB 2521 ACCAATGATGATGATGCTCCAAATGCTGGGTGAGGAGCGGGTCTGATTTGCCGACGAT 2580
QY 2584 CTCGCGCACCGCTTCGCTGCTGACGAGTACAGCGATTTATATCTACTTGAACGACTGGAA 2643
DB 2581 CTCGCGCACCGCTTCGCTGCTGACGAGTACAGCGATTTATATCTACTTGAACGACTGGAA 2640
QY 2644 CTGACAGCATTTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGACCCCTCATCCAGGGC 2703
DB 2641 CTGACAGCATTTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGACCCCTCATCCAGGGC 2700
QY 2704 CACTTGGACTTCGTGATGGACATCTCGTGTGTTCACCTCTCCCGCAGGATGGCTCAAT 2763
DB 2701 CACTTGGACTTCGTGATGGACATCTCGTGTGTTCACCTCTCCCGCAGGATGGCTCAAT 2760
QY 2764 GACTGTATGCAACAACCGGGAGTGTGGGAGCTGTGCTTGCATCCCGGGCGGCAC 2823
DB 2761 GACTGTATGCAACAACCGGGAGTGTGGGAGCTGTGCTTGCATCCCGGGCGGCAC 2820
QY 2824 CGCTGGGCTGGCTCACACTACACCTTGACCCGACGCGCAACTGACAGCCGCGC 2883
DB 2821 CGCTGGGCTGGCTCACACTACACCTTGACCCGACGCGCAACTGACAGCCGCGC 2880
QY 2884 ACCACTTCTTGTCTTGCAGCCAGAAATCTGCCATCACTCGGATGATCCCGGACGACAG 2943
DB 2881 ACCACTTCTTGTCTTGCAGCCAGAAATCTGCCATCACTCGGATGATCCCGGACGACAG 2940
QY 2944 CACAGCCGGATCTCATCTGCTCCCTGATGATGAGGAGAGCTCAAGCCATCGACTAT 3003
DB 2941 CACAGCCGGATCTCATCTGCTCCCTGATGATGAGGAGAGCTCAAGCCATCGACTAT 3000
QY 3004 GACCCACTGGACAACTTTCATCTACTGGGTGGATGGGCGCCAGAACATCAAGCGAGCCAA 3063
DB 3001 GACCCACTGGACAACTTTCATCTACTGGGTGGATGGGCGCCAGAACATCAAGCGAGCCAA 3060
QY 3064 GACGACGGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAAAGGCCAAACCCAGACAG 3123
DB 3061 GACGACGGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAAAGGCCAAACCCAGACAG 3120
QY 3124 CAGCCCAACGACCTCAGCATCGACATCTACAGCCGGAACATGTTCTGACGTTGGAGGCC 3183
DB 3121 CAGCCCAACGACCTCAGCATCGACATCTACAGCCGGAACATGTTCTGACGTTGGAGGCC 3180
QY 3184 ACCAATACATCAACGCTCCACAGGCTGACGGGGAGGCAATGGGGGTGGTCTGCGTGGG 3243
DB 3181 ACCAATACATCAACGCTCCACAGGCTGACGGGGAGGCAATGGGGGTGGTCTGCGTGGG 3240
QY 3244 GACCGCGAACAGCCCAAGGCCATCTGCTCAACCGGAGCGAGGGTACTGTTACTTCAAC 3303
DB 3241 GACCGCGAACAGCCCAAGGCCATCTGCTCAACCGGAGCGAGGGTACTGTTACTTCAAC 3300
QY 3304 AACATGAGGACCGGGACGACAGTCCGACGCGAGCCCTGGAGCGGACCGAGCGCGAG 3363
DB 3301 AACATGAGGACCGGGACGACAGTCCGACGCGAGCCCTGGAGCGGACCGAGCGCGAG 3360
QY 3364 GTCTCTTTCACACCGGCTCTATCCGCCCTGTGGCCCTGTGGTGGACACACACTGGGC 3423

DB 3361 GTCTCTTTCACACCGGCTCTATCCGCCCTGTGGCCCTGTGGTGGACACACACTGGGC 3420
QY 3424 AAGCTGTTCTGGGTGGACGGGACCTGAAAGCGATTTGAGAGCTGTGACCTGTGTCAAGGGCC 3483
DB 3421 AAGCTGTTCTGGGTGGACGGGACCTGAAAGCGATTTGAGAGCTGTGACCTGTGTCAAGGGCC 3480
QY 3484 AACCGCTGACCTTGGAGGAGCGCAACATCGTGAGCCTCTGGGCTGTGACCATCTTGGC 3543
DB 3481 AACCGCTGACCTTGGAGGAGCGCAACATCGTGAGCCTCTGGGCTGTGACCATCTTGGC 3540
QY 3544 AAGCATCTCTACTTGGATCGACCGGACGACAGATGATCGAGCGTGTGGAGAGACCAACC 3603
DB 3541 AAGCATCTCTACTTGGATCGACCGGACGACAGATGATCGAGCGTGTGGAGAGACCAACC 3600
QY 3604 GGGGACAAAGCGGACTCGCATCCAGGGCGGTGTGCGCCACCTCACTGGCATCATGAGTG 3663
DB 3601 GGGGACAAAGCGGACTCGCATCCAGGGCGGTGTGCGCCACCTCACTGGCATCATGAGTG 3660
QY 3664 GAGGAAGTCAAGCTTGGAGGAGTCTCAGCCACCATGTCGCCCTGCAATGTTGGCTGC 3723
DB 3661 GAGGAAGTCAAGCTTGGAGGAGTCTCAGCCACCATGTCGCCCTGCAATGTTGGCTGC 3720
QY 3724 TCCCAACATCTGTATTGTCAAGGGTGTGAGGACACCAAGGTGTGATGTCGCCCTG 3783
DB 3721 TCCCAACATCTGTATTGCCAAGGGTGTGAGGACACCAAGGTGTGATGTCGCCCTG 3780
QY 3784 GTGCTCTCGAGAACTCTGTGACCTGTGGAGAGCGGCCACCTGCTCCCGGACGAGTTT 3843
DB 3781 GTGCTCTCGAGAACTCTGTGACCTGTGGAGAGCGGCCACCTGCTCCCGGACGAGTTT 3840
QY 3844 GCATGTGCCACAGGGAGATCGACTGTATCCCGGGCGCTGGCGCTGTGAGCGCTTCCC 3903
DB 3841 GCATGTGCCACAGGGAGATCGACTGTATCCCGGGCGCTGGCGCTGTGAGCGCTTCCC 3900
QY 3904 GAGTGCATGACACAGAGCGAGAGGGTGTGCCCTGTGCTCCCGGCGGCTTCCC 3963
DB 3901 GAGTGCATGACACAGAGCGAGAGGGTGTGCCCTGTGCTCCCGGCGGCTTCCC 3960
QY 3964 TGCGCGGGGTGAGTGTGGAGCTGCGCTGTGCTGCGAGCGGAGGAGTGTGAG 4023
DB 3961 TGCGCGGGGTGAGTGTGGAGCTGCGCTGTGCTGCGAGCGGAGGAGTGTGAG 4020
QY 4024 GACCGCTCAGACGAGGCGGACTGTGACGCGCATCTGCTGCGCCAAACAGTTCGGGTGTCG 4083
DB 4021 GACCGCTCAGACGAGGCGGACTGTGACGCGCATCTGCTGCGCCAAACAGTTCGGGTGTCG 4080
QY 4084 AGCGGCCAGTGTGCTCTCATCAAAACAGAGTGCAGTCTCTTCCCGGACTGTATGACGGC 4143
DB 4081 AGCGGCCAGTGTGCTCTCATCAAAACAGAGTGCAGTCTCTTCCCGGACTGTATGACGGC 4140
QY 4144 TCCGACGAGCTCATGTGTGAATCAACAGCGCGCTCAGACGACAGCCCGGCCACAGC 4203
DB 4141 TCCGACGAGCTCATGTGTGAATCAACAGCGCGCTCAGACGACAGCCCGGCCACAGC 4200
QY 4204 AGTGCATCGGGCGGCTCATTTGGCATCATCTCTCTCTTCTGTCATGGGTGTCTAT 4263
DB 4201 AGTGCATCGGGCGGCTCATTTGGCATCATCTCTCTCTTCTGTCATGGGTGTCTAT 4260
QY 4264 TTTGTGTGCCAGCGTGTGTCAGCGCTATGCGGGGGCAACCGGGCCCTTCCCGGAC 4323
DB 4261 TTTGTGTGCCAGCGTGTGTCAGCGCTATGCGGGGGCAACCGGGCCCTTCCCGGAC 4320
QY 4324 GAGTATGTACGGGAGCCCGCAGTCCCTCAATTTCTAGCCCGGGGCTTCCCGAG 4383
DB 4321 GAGTATGTACGGGAGCCCGCAGTCCCTCAATTTCTAGCCCGGGGCTTCCCGAG 4380
QY 4384 CATGGCCCTTTCACAGGCACTCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCGTATG 4443
DB 4381 CATGGCCCTTTCACAGGCACTCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCGTATG 4440
QY 4444 GGGGGCGGGGGTGGCCCTGTATCGACCGGAAACGACGTCAAGGGGCTCTGTCAGC 4503

Qy	961	GGCGCTGCTCCACCTGCTGCTGCTGCCAAGCGAGCCTTTCTACATGATGCGCTGC	1020	Db	2060	AAACAGAGTGGCCATCCGCTCAAGGAGGCTCAGCCCTGAGCTTTGAT	2119
Db	980	GGCGCTGCTCCACCTGCTGCTGCTGCCAAGCGAGCCTTTCTACATGATGCGCTGC	1039	Qy	2101	GTGTCCAAACACACATCTACTGACAGACAGCTGAGCTGAAGACATCAGCGCGCTTC	2160
Qy	1021	CCCACGGTGTGACCTGACAGACAAACGGCAGAGCTGTAAAGGCAGAGCCGAGAGGTG	1080	Db	2120	GTGTCCAAACACACATCTACTGACAGACAGCTGAGCTGAAGACATCAGCGCGCTTC	2179
Db	1040	CCCACGGTGTGACCTGACAGACAAACGGCAGAGCTGTAAAGGCAGAGCCGAGAGGTG	1099	Qy	2161	ATGAAACGGGAGCTCGGTGGAGCAGCTGGTGGAGTTGGCTTTGACTACCCCGAGGGCATG	2220
Qy	1081	CTGCTGCTGGCCCGCGGACCGACCTACGGAGGATCTCGCTGGACACGCCGAGCTTCACC	1140	Db	2180	ATGAAACGGGAGCTCGGTGGAGCAGCTGGTGGAGTTGGCTTTGACTACCCCGAGGGCATG	2239
Db	1100	CTGCTGCTGGCCCGCGGACCGACCTACGGAGGATCTCGCTGGACACGCCGAGCTTCACC	1159	Qy	2221	GCCCTTGAATGGATGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGAAATCGAA	2280
Qy	1141	GACATCTGTGCTGACAGTGGACGACATCCGGACCGCAATTCATCGACTACGACCCGCTA	1200	Db	2240	GCCCTTGAATGGATGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGAAATCGAA	2299
Db	1160	GACATCTGTGCTGACAGTGGACGACATCCGGACCGCAATTCATCGACTACGACCCGCTA	1219	Qy	2281	GTGGCGGGCTGGACGGGCGAGTTCGGCAAGTCTCTGCTGGAGGGACTTGGACACCCG	2340
Qy	1201	GAGGGCTATGTCTACTGACAGATGACAGGTGCGGGGCAATCCGACGGGCGTACTGGAC	1260	Db	2300	GTGGCGGGCTGGACGGGCGAGTTCGGCAAGTCTCTGCTGGAGGGACTTGGACACCCG	2359
Db	1220	GAGGGCTATGTCTACTGACAGATGACAGGTGCGGGGCAATCCGACGGGCGTACTGGAC	1279	Qy	2341	AGGTGCTGGCCCTGGATCCCAAGGGCTTACTACTGGACCGAGTGGGGCGGCAAG	2400
Qy	1261	GGGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGTC	1320	Db	2360	AGGTGCTGGCCCTGGATCCCAAGGGCTTACTACTGGACCGAGTGGGGCGGCAAG	2419
Db	1280	GGGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGTC	1339	Qy	2401	CCGAGGATCTGTGCGGGCTTCATGGACGGGACCAACTGCAATGACCTGGTGGACAAGGTG	2460
Qy	1321	GACTGGGTGGCCCGAAACCTCTACTGGACCGACACGGGACCGGACCGCATCGAGGTGACG	1380	Db	2420	CCGAGGATCTGTGCGGGCTTCATGGACGGGACCAACTGCAATGACCTGGTGGACAAGGTG	2479
Db	1340	GACTGGGTGGCCCGAAACCTCTACTGGACCGACACGGGACCGGACCGCATCGAGGTGACG	1399	Qy	2461	GGCCGGGCGCAACGACCTCACCAATGACTACCTGACCAAGCGCTTACTTGGACCGACCTG	2520
Qy	1381	CGCTCAACGGACCTCCCGCAAGATCTCTGTGTGCGAGGACCTTGGACGAGCCCGGAGCC	1440	Db	2480	GGCCGGGCGCAACGACCTCACCAATGACTACCTGACCAAGCGCTTACTTGGACCGACCTG	2539
Db	1400	CGCTCAACGGACCTCCCGCAAGATCTCTGTGTGCGAGGACCTTGGACGAGCCCGGAGCC	1459	Qy	2521	GACACCAACATGATCGAGTCTGTCACCAATGCTGGGTGAGGAGCGGCTGCTGATTCGCCAC	2580
Qy	1441	ATCGCACTGACCCCGTATGAGTGGGCTCATGTGACAGACTGGGAGAGAAACCTTAAA	1500	Db	2540	GACACCAACATGATCGAGTCTGTCACCAATGCTGGGTGAGGAGCGGCTGCTGATTCGCCAC	2599
Db	1460	ATCGCACTGACCCCGTATGAGTGGGCTCATGTGACAGACTGGGAGAGAAACCTTAAA	1519	Qy	2581	GATCTCCGCAACCCGTTCCGCTGTCAGCAGTACAGCGATTAATCTACTTGGACAGACTGG	2640
Qy	1501	ATCGAGTGTGCAACTTTGGATGGGCGAGGCGGGTGTGTGCTCAATGCTCCCTCGGG	1560	Db	2600	GATCTCCGCAACCCGTTCCGCTGTCAGCAGTACAGCGATTAATCTACTTGGACAGACTGG	2659
Db	1520	ATCGAGTGTGCAACTTTGGATGGGCGAGGCGGGTGTGTGCTCAATGCTCCCTCGGG	1579	Qy	2641	AATCTGACAGCATTTGAGCGGGCGGACAGACTAGCGCGCGGAAACCGCACCTCATCCAG	2700
Qy	1561	TGGCCCAACGGCTGGCCCTGGAACCTGACAGAGGGGAGCTCTACTGGGGAGAGCCCAAG	1620	Db	2660	AATCTGACAGCATTTGAGCGGGCGGACAGACTAGCGCGCGGAAACCGCACCTCATCCAG	2719
Db	1580	TGGCCCAACGGCTGGCCCTGGAACCTGACAGAGGGGAGCTCTACTGGGGAGAGCCCAAG	1639	Qy	2701	GGCCACCTGACATTCGTGATGGACATCTCTGTGTTCCACTCTCCCGCAGGATGGCCTC	2760
Qy	1621	ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAAGGCGGACCCCTCTCGAGGAC	1680	Db	2720	GGCCACCTGACATTCGTGATGGACATCTCTGTGTTCCACTCTCTCCCGCAGGATGGCCTC	2779
Db	1640	ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAAGGCGGACCCCTCTCGAGGAC	1699	Qy	2761	AATGACTGTATGCAACAAACGGGACGTGGGACGTGTGCTTGGCCATCCCGCGGCGC	2820
Qy	1681	AAGCTCCCGCACATTTTCGGGTTTCAACGCTGCTGGGGGACTTTCATCTACTGAGCTGACTGG	1740	Db	2780	AATGACTGTATGCAACAAACGGGACGTGTGGGACGTGTGCTTGGCCATCCCGCGGCGC	2839
Db	1700	AAGCTCCCGCACATTTTCGGGTTTCAACGCTGCTGGGGGACTTTCATCTACTGAGCTGACTGG	1759	Qy	2821	CACCGCTGGCGCTGGCCCTCACACTACCCCTGGACCCCGCAGCAGCGCACTGACGCCG	2880
Qy	1741	CAGCGCGCAGCATCGAGCGGGTGCACAGGTCAAGGCGAGCGGAGCTCATCATTTGAC	1800	Db	2840	CACCGCTGGCGCTGGCCCTCACACTACCCCTGGACCCCGCAGCAGCGCACTGACGCCG	2899
Db	1760	CAGCGCGCAGCATCGAGCGGGTGCACAGGTCAAGGCGAGCGGAGCTCATCATTTGAC	1819	Qy	2881	CCACACCTTCTTGTGTTTTCAGCCAGAAATCTCCCATCAGTCGAGTGTATCCCGAGCCAC	2940
Qy	1801	CAGCTGCCCGACCTGATGGGCTCAAGCTGCTGATGTGGCCAAAGTCTCGGAAACCAAC	1860	Db	2900	CCACACCTTCTTGTGTTTTCAGCCAGAAATCTCCCATCAGTCGAGTGTATCCCGAGCCAC	2959
Db	1820	CAGCTGCCCGACCTGATGGGCTCAAGCTGCTGATGTGGCCAAAGTCTCGGAAACCAAC	1879	Qy	2941	CAGCACAGCCCGGATCTCATCTCTGCCCTTGCATGGAATGAGGAGAACTGAAAGCCATCGAC	3000
Qy	1861	CCGTGTGCGGACAGGAAACGGGGGTGACGCCACCTGTCTTTCACACCCCGACCAACC	1920	Db	2960	CAGCACAGCCCGGATCTCATCTCTGCCCTTGCATGGAATGAGGAGAACTGAAAGCCATCGAC	3019
Db	1880	CCGTGTGCGGACAGGAAACGGGGGTGACGCCACCTGTCTTTCACACCCCGACCAACC	1939	Qy	3001	TATGACCCACTGGACAGATTCATCTACTGGTGGATGGGGCGGACAGAACTCAAGCGAGCC	3060
Qy	1921	CGGTGTGCTGCCCATCGGCTGGAGCTGTGAGTGACATGAAGAAGCTGATCGTGCCT	1980	Db	3020	TATGACCCACTGGACAGATTCATCTACTGGTGGATGGGGCGGACAGAACTCAAGCGAGCC	3079
Db	1940	CGGTGTGCTGCCCATCGGCTGGAGCTGTGAGTGACATGAAGAAGCTGATCGTGCCT	1999	Qy	3061	AAGGACGACGGGACCCAGCCCTTGTGTTTTCAGCTCTCTGTGAGCCAAAGGCAAAACCCAGAC	3120
Qy	1981	GAGGCTTCTTGTGTTTTCACAGCAGAGCCCGCCATCCACAGGATCTCCCTCGAGACCAAT	2040	Db	3080	AAGGACGACGGGACCCAGCCCTTGTGTTTTCAGCTCTCTGTGAGCCAAAGGCAAAACCCAGAC	3139
Db	2000	GAGGCTTCTTGTGTTTTCACAGCAGAGCCCGCCATCCACAGGATCTCCCTCGAGACCAAT	2059	Qy	3121	AGGACGACCCCAACGACCTCAGCATCGACATCTTACAGCCGAGACACTGTTCTTGGACGTGGAG	3180
Qy	2041	AAACACGACGTGGCCATCCCGCTCAACGGGCGTCAAGGAGGCGCTCAGCGCTTGTGAT	2100				

[illegible]

Qy	4261	TATTTTGTGTGCCAGCGCGTGTGTGTGCCAGCGCTATGCGGGGCGCAACGCGGCCCTTCCCG	4320
Db	4280	TATTTTGTGTGCCAGCGCGTGTGTGTGCCAGCGCTATGCGGGGCGCAACGCGGCCCTTCCCG	4339
Qy	4321	CACGAGTATGTACGCGGGACCCGCGACGTGCCCTCAATTTTCAATAGCCCGGCGGTTC	4380
Db	4340	CACGAGTATGTACGCGGGACCCGCGACGTGCCCTCAATTTTCAATAGCCCGGCGGTTC	4399
Qy	4381	CAGCATGGCCCCCTTTCACAGGCATGCGATGCGGAAAGTCCATGATGAGCTCCGTGAGCCTG	4440
Db	4400	CAGCATGGCCCCCTTTCACAGGCATGCGATGCGGAAAGTCCATGATGAGCTCCGTGAGCCTG	4459
Qy	4441	ATGGGGGGCGGGCGGGGTGCCCTGTACGACGGAAACCACTGTCACAGGGGCTGTGTC	4500
Db	4460	ATGGGGGGCGGGCGGGGTGCCCTGTACGACGGAAACCACTGTCACAGGGGCTGTGTC	4519
Qy	4501	AGCAGCTCGTCCAGCAGCAAGAGGCACGCTGTACCGCGCATCTGAAACCGCGCCCTCC	4560
Db	4520	AGCAGCTCGTCCAGCAGCAAGAGGCACGCTGTACCGCGCATCTGAAACCGCGCCCTCC	4579
Qy	4561	CCGCCCCAGCACCCCTCCCTGTACCAATGAGCATGTTCTTACTTTCAACATTTCCGGCC	4620
Db	4580	CCGCCCCAGCACCCCTCCCTGTACCAATGAGCATGTTCTTACTTTCAACATTTCCGGCC	4639
Qy	4621	ACTCGAGACGGTACAGGCCCTACATCATTCGAGGAAATGGCGCCCCGAGCAGCGCCCTGC	4680
Db	4640	ACTCGAGACGGTACAGGCCCTACATCATTCGAGGAAATGGCGCCCCGAGCAGCGCCCTGC	4699
Qy	4681	AGCACCGACGTGTGTACACGCACTACAGCGCCAGCGCTGTGAAGGACGACAAAGTACTAC	4740
Db	4700	AGCACCGACGTGTGTACACGCACTACAGCGCCAGCGCTGTGAAGGACGACAAAGTACTAC	4759
Qy	4741	CTGGAATTGAATCTCGGACTCAGACCGCTATPCACCCCCACCCACGCCCCACAGCCAGTAC	4800
Db	4760	CTGGAATTGAATCTCGGACTCAGACCGCTATPCACCCCCACCCACGCCCCACAGCCAGTAC	4819
Qy	4801	CTGTGCGCGAGACAGCTGCCCGCCCTGCCCGCCACCGAGAGGAGCTACTTCCATCTC	4860
Db	4820	CTGTGCGCGAGACAGCTGCCCGCCCTGCCCGCCACCGAGAGGAGCTACTTCCATCTC	4879
Qy	4861	TTCCCGGCCCTCCGTCCCCCTGCACGGACTCATCTGACCTCGCGCGGGCCCACTCTGGC	4920
Db	4880	TTCCCGGCCCTCCGTCCCCCTGCACGGACTCATCTGACCTCGCGCGGGCCCACTCTGGC	4939
Qy	4921	TTCTCTGTGCCCTGTAAATAGTTTTAAATATGAAACAAGAAAAAATATATTTTATGAT	4980
Db	4940	TTCTCTGTGCCCTGTAAATAGTTTTAAATATGAAACAAGAAAAAATATATTTTATGAT	4999
Qy	4981	TTAAAAATAAATATATTTGGGATTTTAAAAACATGAGAAATGTGAACTGTGATGGGTG	5040
Db	5000	TTAAAAATAAATATATTTGGGATTTTAAAAACATGAGAAATGTGAACTGTGATGGGTG	5059
Qy	5041	GGCAGGCGTCGGAGAACTTTGTA	5063
Db	5060	GGCAGGCGTCGGAGAACTTTGTA	5082

RESULT 15
ADB99039
ID ADB99039 standard: DNA: 5120 BP.

AA ADB99039;

DT 04-DEC-2003 (first entry)

XX DE LRP5 variant coding sequence.

XX Osteopathic: Gene therapy: High Bone Mass: HBM: LRP5: Zmax1: LRP6: KW

KW bone mass modulation; osteoporosis; gene; mutein; mutant; ds.
XY

OS Synthetic.

QY	1561	TGGCCCAAGCGGCTGGCCCTGGACCTGACGAGGGGAGCTTACTTGGGGAGACCCGAAG	1620	Db	2660	AATCTGCACAGCATTTAGCGGGCCGACAAGACTAGCGCCGGAAACCGCACCCCTCATCCAG	2719
Db	1580	TGGCCCAAGCGGCTGGCCCTGGACCTGACGAGGGGAGAGCTTACTTGGGGAGACCCGAAG	1639	QY	2701	GGCCACCTGGAGCTTCGTGATGGACATCTCTGTGTTCACATCTCTCCGCGCAGGATGGCCTC	2760
QY	1621	ACAGACAAGATCGAGGTGATCAATTTGATGGGACGAAGAGCGGACCCCTCTCTGGAGAC	1680	Db	2720	GGCCACCTGGAGCTTCGTGATGGACATCTCTGTGTTCACATCTCTCCGCGCAGGATGGCCTC	2779
Db	1640	ACAGACAAGATCGAGGTGATCAATTTGATGGGACGAAGAGCGGACCCCTCTCTGGAGAC	1699	QY	2761	AATGACTGTATGCAACAACGCGGAGGTGTGGGAGCTGTGCTTGGCCATCCCCGGCGGC	2820
QY	1681	AAGCTCCGCGCAATTTTGGGGTTACGCTGTCTGGGGACTTCACTACTATGAGACTGACTGG	1740	Db	2780	AATGACTGTATGCAACAACGCGGAGGTGTGGGAGCTGTGCTTGGCCATCCCCGGCGGC	2839
Db	1700	AAGCTCCGCGCAATTTTGGGGTTACGCTGTCTGGGGACTTCACTACTATGAGACTGACTGG	1759	QY	2821	CACCGCTCGGCTGGCTCACAATCTACACCTCTGGAACCCAGCAGCCGCAATGACGACCG	2880
QY	1741	CAGCCCGCGAGCATCGAGCGGTGACAAAGTCAAGGTCAGGCGGACCGCGGAGCTCATATTGAC	1800	Db	2840	CACCGCTCGGCTGGCTCACAATCTACACCTCTGGAACCCAGCAGCCGCAATGACGACCG	2899
Db	1760	CAGCCCGCGAGCATCGAGCGGTGACAAAGTCAAGGTCAGGCGGACCGCGGAGCTCATATTGAC	1819	QY	2881	CCCACCACTTCTTGTCTTTCAGCCAGAAATCTGCCATCAGTCGGATGATCCCGGACGAC	2940
QY	1801	CAGCTGCCCGACCTGATGGGGCTCAAGAGCTGTGAATGTGGCCAAAGGTGTCGGAACCAAC	1860	Db	2900	CCCACCACTTCTTGTCTTTCAGCCAGAAATCTGCCATCAGTCGGATGATCCCGGACGAC	2959
Db	1820	CAGCTGCCCGACCTGATGGGGCTCAAGAGCTGTGAATGTGGCCAAAGGTGTCGGAACCAAC	1879	QY	2941	CAGCAGAGCCCGGATCTCATCTCTGCCCCCTGTGATGGAATGGAACGCTCAAAAGCCATCGAC	3000
QY	1861	CGTGTGGGACAGAAAGGGGGGTGACGCCACCTGTGTGTTCTTCAACCCCAACCAACC	1920	Db	2960	CAGCAGAGCCCGGATCTCATCTCTGCCCCCTGTGATGGAATGGAACGCTCAAAAGCCATCGAC	3019
Db	1880	CGTGTGGGACAGAAAGGGGGGTGACGCCACCTGTGTGTTCTTCAACCCCAACCAACC	1939	QY	3001	TATGACCCACTGGACAAAGTTTCTACTTGGGTGGATGGCGCCAGAAATCAAGCGAGCC	3060
QY	1921	CGTGTGGCTGCCCATCGGCTGGAGCTGTGAGTGCATGAGACCTGTCATCGTGCCT	1980	Db	3020	TATGACCCACTGGACAAAGTTTCTACTTGGGTGGATGGCGCCAGAAATCAAGCGAGCC	3079
Db	1940	CGTGTGGCTGCCCATCGGCTGGAGCTGTGAGTGCATGAGACCTGTCATCGTGCCT	1999	QY	3061	AAGGACGACGGGACCCAGCCCTTTTGTGACCTCTCTGAGCCAGGCGCAAAACCCAGAC	3120
QY	1981	GAGGCTTCTTGTGTTTCAACGACGAGCGCGCATCCACAGGATCTCCCTCGAGACCAAT	2040	Db	3080	AAGGACGAGGGGACCCAGCCCTTTTGTGACCTCTCTGAGCCAGGCGCAAAACCCAGAC	3139
Db	2000	GAGGCTTCTTGTGTTTCAACGACGAGCGCGCATCCACAGGATCTCCCTCGAGACCAAT	2059	QY	3121	AGGACAGCCCGGACCTCAGCATCGACATCTACAGCCGGAACCTGTCTTGACAGTGCAG	3180
QY	2041	AACACGAGTGGCCATCCCGCTCAGCGGGGTCAAGGAGGCTCAGCCCTGGAATTTGAT	2100	Db	3140	AGGACAGCCCGGACCTCAGCATCGACATCTACAGCCGGAACCTGTCTTGACAGTGCAG	3199
Db	2060	AACACGAGTGGCCATCCCGCTCAGCGGGGTCAAGGAGGCTCAGCCCTGGAATTTGAT	2119	QY	3181	GCCACCAATACCATCAAGCTCCAGGCTGAGCGGGGAGCCATGCGGGGTGCTGCTGCT	3240
QY	2101	GTGTCCAAACCAACATCTACTGGACAGACGCTCAGCCTGAAGACCATCAGCGCGGCTTC	2160	Db	3200	GCCACCAATACCATCAAGCTCCAGGCTGAGCGGGGAGCCATGCGGGGTGCTGCTGCT	3259
Db	2120	GTGTCCAAACCAACATCTACTGGACAGACGCTCAGCCTGAAGACCATCAGCGCGGCTTC	2179	QY	3241	GGGACCGCGCAAGCCCGAGGGCCATCGTGTCAACCGGGAGCGAGGTACTCTGTACTTC	3300
QY	2161	ATGAACGGGAGCTCGGTGGAGCACTGTGTGAGTTTGGCTTGAATACCCCGAGGGCATG	2220	Db	3260	GGGACCGCGCAAGCCCGAGGGCCATCGTGTCAACCGGGAGCGAGGTACTCTGTACTTC	3319
Db	2180	ATGAACGGGAGCTCGGTGGAGCACTGTGTGAGTTTGGCTTGAATACCCCGAGGGCATG	2239	QY	3301	ACCAACATGAGGACCGGGGAGCCAAAGTCGAAACCGGAGCCCTTGGACGCGCACCGAGCGC	3360
QY	2221	CCCGTTGACTGGATGGGCAAGAACTCTACTTGGGCGGACACTGGGACCAACAGAAATCGAA	2280	Db	3320	ACCAACATGAGGACCGGGGAGCCAAAGTCGAAACCGGAGCCCTTGGACGCGCACCGAGCGC	3379
Db	2240	CCCGTTGACTGGATGGGCAAGAACTCTACTTGGGCGGACACTGGGACCAACAGAAATCGAA	2299	QY	3361	GAGTCTCTTTCACACCGGCTCATCCGCTCTGTGCGCTGTGTGGTGGACAAACACACTG	3420
QY	2281	GTGGCGGCTGGACGGGCACTTCGGCAAGTCTCTGCTGGAGGACCTTGGACAAACCGG	2340	Db	3380	GAGTCTCTTTCACACCGGCTCATCCGCTCTGTGCGCTGTGTGGTGGACAAACACACTG	3439
Db	2300	GTGGCGGCTGGACGGGCACTTCGGCAAGTCTCTGCTGGAGGACCTTGGACAAACCGG	2359	QY	3421	GGCAAGCTTGTGGGTGGACCGGCACTCTGAAAGCCGCAATGAGAGCTGTGACCTGTGAGG	3480
QY	2341	AGTGTGCTGGCTGGATFCCCAACCAAGGGCTACATCTACTGGAACGAGTGGGGCGGCAAG	2400	Db	3440	GGCAAGCTTGTGGGTGGACCGGCACTCTGAAAGCCGCAATGAGAGCTGTGACCTGTGAGG	3499
Db	2360	AGTGTGCTGGCTGGATFCCCAACCAAGGGCTACATCTACTGGAACGAGTGGGGCGGCAAG	2419	QY	3481	GCCAAACCGCTGACCTTGGAGGACCGCAACATCGTGCAGCTCTGGGCTGACCATCCTT	3540
QY	2401	CCGAGGATCGGCGGCTTCTATGAGCGGACCAACTGCATGACGCTGTGTGGAACAGGTG	2460	Db	3500	GCCAAACCGCTGACCTTGGAGGACCGCAACATCGTGCAGCTCTGGGCTGACCATCCTT	3559
Db	2420	CCGAGGATCGGCGGCTTCTATGAGCGGACCAACTGCATGACGCTGTGTGGAACAGGTG	2479	QY	3541	GGCAAGCATCTCTACTTGGATCGACCGGACGAGCAGATGATCGAGCGGTGTGGAAGAGAC	3600
QY	2461	GGCGGGGCAACGACCTCACAATTGACTACGCTGACACGCGCTCTACTTGGACCGACCTG	2520	Db	3560	GGCAAGCATCTCTACTTGGATCGACCGGACGAGCAGATGATCGAGCGGTGTGGAAGAGAC	3619
Db	2480	GGCGGGGCAACGACCTCACAATTGACTACGCTGACACGCGCTCTACTTGGACCGACCTG	2539	QY	3601	ACCGGGGACAGCGGATCTCGCATCCAGGCGGTGTGCGCCACCTCTACTTGGATCGATGCA	3660
QY	2521	GACACCAACATGATGAGTCTGCAACATGCTGGGTGAGGAGCGGGTCTGATTTGCCGAC	2580	Db	3620	ACCGGGGACAGCGGATCTCGCATCCAGGCGGTGTGCGCCACCTCTACTTGGATCGATGCA	3679
Db	2540	GACACCAACATGATGAGTCTGCAACATGCTGGGTGAGGAGCGGGTCTGATTTGCCGAC	2599	QY	3661	GTGGAGGAAGTTCAGCCTGGAGGAGTTCTCAGCCCCACCATGTGCGCCGTGACAAATGGTGGC	3720
QY	2581	GATCTCCCGCACCGCTTCCGCTCAGCGCAGTACACCGGATTTATCTACTTGCACAGACTGG	2640	Db	3680	GTGGAGGAAGTTCAGCCTGGAGGAGTTCTCAGCCCCACCATGTGCGCCGTGACAAATGGTGGC	3739
Db	2600	GATCTCCCGCACCGCTTCCGCTCAGCGCAGTACACCGGATTTATCTACTTGCACAGACTGG	2659	QY	3721	TGCTCCCAACATCTGTATTGCGAAAGGTGATGGGACCAACGCGTCTCATGCCCCAGTCCAC	3780

THIS PAGE BLANK (USPTO)

Db 2480 GCGCGGCGCAACGACCTCACATTGACTACGCTGACACGAGCGCTCTACTGACCGACCGT 2539
QY 2521 GACACCAACATGATCGAGTCTGTCACCAATGCTGGGTGAGGAGCGGTGCTGATTGGCGAC 2580
Db 2540 GACACCAACATGATCGAGTCTGTCACCAATGCTGGGTGAGGAGCGGTGCTGATTGGCGAC 2599
QY 2581 GATCTCCCGCAGCCCGTTCGGTCTGACGAGTACAGCGATTATATCTACTGGACAGACTGG 2640
Db 2600 GATCTCCCGCAGCCCGTTCGGTCTGACGAGTACAGCGATTATATCTACTGGACAGACTGG 2659
QY 2641 AATCTGCAACAGATTGAGCGGCGGCAAAAGACTAGCGGCGGAAACCGCACCTCATCCAG 2700
Db 2660 AATCTGCAACAGATTGAGCGGCGGCAAAAGACTAGCGGCGGAAACCGCACCTCATCCAG 2719
QY 2701 GGCACCTGGACTTCTGATGAGCAATCTGCTGTTCCACTCTCCCGCAGGATGGCCCTC 2760
Db 2720 GGCACCTGGACTTCTGATGAGCAATCTGCTGTTCCACTCTCCCGCAGGATGGCCCTC 2779
QY 2761 AATGACTGTATGCAAAACAAAGCGGAGTGTGGGAGCTGTGCTTGGCCATCCCGGGCGC 2820
Db 2780 AATGACTGTATGCAAAACAAAGCGGAGTGTGGGAGCTGTGCTTGGCCATCCCGGGCGC 2839
QY 2821 CACCGCTCGGCTGCGCTCACTACATACCTGGAACCCGAGCGGCAACTGACGCGC 2880
Db 2840 CACCGCTCGGCTGCGCTCACTACATACCTGGAACCCGAGCGGCAACTGACGCGC 2899
QY 2881 CCCACCACTTCTGCTGTTGAGCCAGAAATCTGCCATCAGTCGGATATCCCGAGCAG 2940
Db 2900 CCCACCACTTCTGCTGTTGAGCCAGAAATCTGCCATCAGTCGGATATCCCGAGCAG 2959
QY 2941 CAGCACAGCCCGGATCTCATCTGCGCTGATGGAGCTGAGGAAAGCTCAAGCCATCGAC 3000
Db 2960 CAGCACAGCCCGGATCTCATCTGCGCTGATGGAGCTGAGGAAAGCTCAAGCCATCGAC 3019
QY 3001 TATGACCACTGGAAGTTTCACTACTGCTGATGGGCGGCCAGAAACATCAAGCGAGCC 3060
Db 3020 TATGACCACTGGAAGTTTCACTACTGCTGATGGGCGGCCAGAAACATCAAGCGAGCC 3079
QY 3061 AAGGACGAGGAGCCCGCTTGTGTTGACCTCTCTGAGGAGGCGGCAAAACCCAGAC 3120
Db 3080 AAGGACGAGGAGCCCGCTTGTGTTGACCTCTCTGAGGAGGCGGCAAAACCCAGAC 3139
QY 3121 AGGACAGCCCGACGCTCAGCATCGACATCTACAGCGGAGCACTGTTCTGAGAGTGCAG 3180
Db 3140 AGGACAGCCCGACGCTCAGCATCGACATCTACAGCGGAGCACTGTTCTGAGAGTGCAG 3199
QY 3181 GCCACCAATACATCAACGCTCCACAGGCTGAGCGGGAGCCATGGGGGTGGTGTGGT 3240
Db 3200 GCCACCAATACATCAACGCTCCACAGGCTGAGCGGGAGCCATGGGGGTGGTGTGGT 3259
QY 3241 GGGGACCGGACAGGCCCGAGGCCATCTGCTCAACGGGAGCGAGGTTACTGTACTTC 3300
Db 3260 GGGGACCGGACAGGCCCGAGGCCATCTGCTCAACGGGAGCGAGGTTACTGTACTTC 3319
QY 3301 ACCAATGTCAGGACCGGCGAGCCAAAGATCGAACCGCGAGCCCTGAGCGGACCGAGCGC 3360
Db 3320 ACCAATGTCAGGACCGGCGAGCCAAAGATCGAACCGCGAGCCCTGAGCGGACCGAGCGC 3379
QY 3361 GAGGTCTCTTCAACACCGGCTCATCCGCTGTTGGCCCTGGTGGTGGACAAACACTG 3420
Db 3380 GAGGTCTCTTCAACACCGGCTCATCCGCTGTTGGCCCTGGTGGTGGACAAACACTG 3439
QY 3421 GGCAGCTGTTCTGGGTGGACCGGACCTGAGCGCATTTGAGAGCTGTGACTGTGAGG 3480
Db 3440 GGCAGCTGTTCTGGGTGGACCGGACCTGAGCGCATTTGAGAGCTGTGACTGTGAGG 3499
QY 3481 GCCAACCGCTGACCTGAGGAGCGCAACATCTGTCAGGCTCTGGGCTGACCATCCTT 3540
Db 3500 GCCAACCGCTGACCTGAGGAGCGCAACATCTGTCAGGCTCTGGGCTGACCATCCTT 3559
QY 3541 GGCAGCACTCTACTGTGATCGACCGGACGAGCATGATCGAGCGTGTGGAGAGACC 3600

Db 3560 GGCAGCATCTCTACTGATCGACCGCGCAGCAGCATGATCGAGCGGTGTGGAGAGACC 3619
QY 3601 ACCGGGAGCAAGCGGACTCGCATCCAGGGCGGTGCGCCACCTCCTACTGGCATCCATGCA 3660
Db 3620 ACCGGGAGCAAGCGGACTCGCATCCAGGGCGGTGCGCCACCTCCTACTGGCATCCATGCA 3679
QY 3661 GTGAGGAAGTCAAGCTCGGAGGAGTCTTCAAGCCCAACCATGTGCGCGTGACAAATGGTGGC 3720
Db 3680 GTGAGGAAGTCAAGCTCGGAGGAGTCTTCAAGCCCAACCATGTGCGCGTGACAAATGGTGGC 3739
QY 3721 TGTCTCCACATCTCTGATTTGCCAAGGTGATGGGACACACAGGTGCTATGCGCAGTCCAC 3780
Db 3740 TGTCTCCACATCTCTGATTTGCCAAGGTGATGGGACACACAGGTGCTATGCGCAGTCCAC 3799
QY 3781 CTGCTGTCTCTGCAAGAACTCTGCTGACCTGTGAGAGCGCGCCCACTGCTCCCGGACGAG 3840
Db 3800 CTGCTGTCTCTGCAAGAACTCTGCTGACCTGTGAGAGCGCGCCCACTGCTCCCGGACGAG 3859
QY 3841 TTTGATGTGTCACAGGGGAGATCGACTGTATCCCGGGGCTGTGGCTGTGAACGGCTTT 3900
Db 3860 TTTGATGTGTCACAGGGGAGATCGACTGTATCCCGGGGCTGTGGCTGTGAACGGCTTT 3919
QY 3901 CCCAGTGTGATGACAGAGCGAGAGGGCTGCGCCGCTGCTCCCGCGCCAGTTTC 3960
Db 3920 CCCAGTGTGATGACAGAGCGAGAGGGCTGCGCCGCTGCTCCCGCGCCAGTTTC 3979
QY 3961 CCTGCGCGGGGTGAGTGTGAGCCTGCGCTGCGCTGCGAGCGGAGGAGGACTGT 4020
Db 3980 CCTGCGCGGGGTGAGTGTGAGCCTGCGCTGCGCTGCGAGCGGAGGAGGACTGT 4039
QY 4021 CAGACCGCTCAGACAGGGGAGTGTGACCGCCATCTGCTTGCCTGCGGAGGAGGACTGT 4080
Db 4040 CAGACCGCTCAGACAGGGTGTGAGTGTGAGCCTGCTGCGCTGCGAGCGGAGGAGGACTGT 4099
QY 4081 GCGAGCGGCGAGTGTGCTCATCAACAGAGTGTGAGTCTCTCTCTGCTCATGGTGGTGTTC 4140
Db 4100 GCGAGCGGCGAGTGTGCTCATCAACAGAGTGTGAGTCTCTCTCTGCTCATGGTGGTGTTC 4159
QY 4141 GGTCTCCAGAGCTCATGTGTGAATCAACAAAGCGGCTCAGACGAGCAGCGCGGCCAC 4200
Db 4160 GGTCTCCAGAGCTCATGTGTGAATCAACAAAGCGGCTCAGACGAGCAGCGCGGCCAC 4219
QY 4201 AGCAGTGTGCTCGGCGCGCTCATTTGGCATCATCTCTCTCTCTGCTCATGGTGGTGTTC 4260
Db 4220 AGCAGTGTGCTCGGCGCGCTCATTTGGCATCATCTCTCTCTCTGCTCATGGTGGTGTTC 4279
QY 4261 TATTTTGTGTCAGAGCGGTGTGCGAGCGTATGCGGGGGCCAAAGCGGCGCTTCGCG 4320
Db 4280 TATTTTGTGTCAGAGCGGTGTGCGAGCGTATGCGGGGGCCAAAGCGGCGCTTCGCG 4339
QY 4321 CACAGTGTGTCAGGGGAGCCCGCAGTGTGCGCTCAATTTTCATAGCCCGGGGGTTC 4380
Db 4340 CACAGTGTGTCAGGGGAGCCCGCAGTGTGCGCTCAATTTTCATAGCCCGGGGGTTC 4399
QY 4381 CAGCATGGCCCTTTCACAGGCGATCGCATGCGGAAAGTCCATGATGAGTCTCGTGAAGCTG 4440
Db 4400 CAGCATGGCCCTTTCACAGGCGATCGCATGCGGAAAGTCCATGATGAGTCTCGTGAAGCTG 4459
QY 4441 ATGGGGGGCGGGGGGGTGTGCGAGCGTATGCGGGGGCCAAAGCGGCGCTTCGCG 4500
Db 4460 ATGGGGGGCGGGGGGGTGTGCGAGCGTCTTACAGACCGGAAACAGCTCAAGGGGCTTCGTC 4519
QY 4501 AGCAGTGTGTCAGAGCGGAGCCCGCAGTGTGACCGCGATCTGAACCCCGCGGCTTC 4560
Db 4520 AGCAGTGTGTCAGAGCGGAGCCCGCAGTGTGACCGCGATCTGAACCCCGCGGCTTC 4579
QY 4561 CCGGCCACGAGCCCTCTCTGTACAAATGAGCATGTTTCTACTCTTCAAACATTCGCGCC 4620
Db 4580 CCGGCCACGAGCCCTCTCTGTACAAATGAGCATGTTTCTACTCTTCAAACATTCGCGCC 4639
QY 4621 ACTCGGAGACGTTACAGGCGCTCATCATTCGAGGAATGGCGGCCCGCCAGACGCGCTTC 4680
Db 4640 ACTCGGAGACGTTACAGGCGCTCATCATTCGAGGAATGGCGGCCCGCCAGACGCGCTTC 4699

Db 1280 GGGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGTCT 1339
QY 1321 GACTGGGTGGCCCGAAACCTCTACTGGACCGACAGCGGCGACGCGGATCGAGTGGACG 1380
Db 1340 GACTGGGTGGCCCGAAACCTCTACTGGACCGACAGCGGCGACGCGGATCGAGTGGACG 1399
QY 1381 CGCCTCAACGGGACCTCCCGCAAGATCTCTGGTGTGGAGGACCTGGAGAGCCCGAGGCC 1440
Db 1400 CGCCTCAACGGGACCTCCCGCAAGATCTCTGGTGTGGAGGACCTGGAGAGCCCGAGGCC 1459
QY 1441 ATCGACATGACACCCGATGATGGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAA 1500
Db 1460 ATCGACATGACACCCGATGATGGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAA 1519
QY 1501 ATCGAGTGTGCCAATCTTGATGGGAGGAGGAGCGGCTGTGCTGATCAATGCCCTCCCTCGG 1560
Db 1520 ATCGAGTGTGCCAATCTTGATGGGAGGAGGAGCGGCTGTGCTGATCAATGCCCTCCCTCGG 1579
QY 1561 TGGCCCAACGGGCTGGCCCTGGACCTGACGAGGAGGAGGCTTACTGGGGAGAGGCCAAG 1620
Db 1580 TGGCCCAACGGGCTGGCCCTGGACCTGACGAGGAGGAGGCTTACTGGGGAGAGGCCAAG 1639
QY 1621 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAGAGGCGGACCTCTCGAGGAC 1680
Db 1640 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAGAGGCGGACCTCTCGAGGAC 1699
QY 1681 AAGCTCCCGCACATTTTCGGGTTTACGCTGTCTGGGGGACTTCTACTGTGACACTGCTG 1740
Db 1700 AAGCTCCCGCACATTTTCGGGTTTACGCTGTCTGGGGGACTTCTACTGTGACACTGCTG 1759
QY 1741 CAGCGCGCAGCATCGAGCGGTGCACAAAGTCAAGGCCAGCGGAGCTCATCATTTGAC 1800
Db 1760 CAGCGCGCAGCATCGAGCGGTGCACAAAGTCAAGGCCAGCGGAGCTCATCATTTGAC 1819
QY 1801 CAGCTGCCCGACCTGATGGGCTCAAAAGCTGTGAATGGCCAAAGTGTGGCAAGCTTCA 1860
Db 1820 CAGCTGCCCGACCTGATGGGCTCAAAAGCTGTGAATGGCCAAAGTGTGGCAAGCTTCA 1879
QY 1861 CCGTGTGGGACAGGACCGGGGTGCAGCCACCTGTGCTTCTTCAACCCCGACGCAAC 1920
Db 1880 CCGTGTGGGACAGGACCGGGGTGCAGCCACCTGTGCTTCTTCAACCCCGACGCAAC 1939
QY 1921 CCGTGTGGCTCCCGCATCGGCTCGAGCTGTGAGTGACATGAAGACCTGTGCTGCT 1980
Db 1940 CCGTGTGGCTCCCGCATCGGCTCGAGCTGTGAGTGACATGAAGACCTGTGCTGCT 1999
QY 1981 GAGGCTTCTTGGTCTTCAACAGCAGAGCGGCCATCCACAGGATCTTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTTCTTGGTCTTCAACAGCAGAGCGGCCATCCACAGGATCTTCCCTCGAGACCAAT 2059
QY 2041 AACCAACGCTGGCCATCCCGCTCACGGGCGTCAAGGAGGCTCAGCGCTGGACTTTGAT 2100
Db 2060 AACCAACGCTGGCCATCCCGCTCACGGGCGTCAAGGAGGCTCAGCGCTGGACTTTGAT 2119
QY 2101 GTGTCCAAACACACATCTACTGGACAGAGCTCAGCCTGAAGACCATCAGCGCGCTTC 2160
Db 2120 GTGTCCAAACACACATCTACTGGACAGAGCTCAGCCTGAAGACCATCAGCGCGCTTC 2179
QY 2161 ATGAACCGGAGCTCGGTGGACACGTGGTGGAGTTTGGCCCTTGTACTACCCCGAGGGCATG 2220
Db 2180 ATGAACCGGAGCTCGGTGGACACGTGGTGGAGTTTGGCCCTTGTACTACCCCGAGGGCATG 2239
QY 2221 GCCCTTGAATGGGACAGAACCTTACTTGGGCGGACACTGGGACCAACAGATCGAA 2280
Db 2240 GCCCTTGAATGGGACAGAACCTTACTTGGGCGGACACTGGGACCAACAGATCGAA 2299
QY 2281 GTGGCGGGCTGGACGGGACGTTCGGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCG 2340
Db 2300 GTGGCGGGCTGGACGGGACGTTCGGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCG 2359
QY 2341 AGGTGCTGGCCCTGGATCCCAACAAAGGCTTACATCTACTGGACCGGAGTGGGGCGCAAG 2400

Db 2360 AGGTGCTGGCCCTGGATCCCAACAAAGGCTTACTACTGACCGAGTGGGGCGGCAAG 2419
QY 2401 CCGAGGATCGTGGGGCTTTCATGACGGGACCAACTGCATGACGCTGGTGGACAGGTG 2460
Db 2420 CCGAGGATCGTGGGGCTTTCATGACGGGACCAACTGCATGACGCTGGTGGACAGGTG 2479
QY 2461 GCGCGGGCCAAACGACCTCACCATTTGACTACGCTGACAGCGCTTACTTGGACCGACCTG 2520
Db 2480 GCGCGGGCCAAACGACCTCACCATTTGACTACGCTGACAGCGCTTACTTGGACCGACCTG 2539
QY 2521 GACACCAACATGATCGAGTCTCTCAACATGCTGGGTGAGGAGCGGTCTGTGATGCCGAC 2580
Db 2540 GACACCAACATGATCGAGTCTCTCAACATGCTGGGTGAGGAGCGGTCTGTGATGCCGAC 2599
QY 2581 GATCTCCCGCACCGCTTCGGTCTGACGAGTACAGCGGATATATCTACTTGGACAGACTGG 2640
Db 2600 GATCTCCCGCACCGCTTCGGTCTGACGAGTACAGCGGATATATCTACTTGGACAGACTGG 2659
QY 2641 AATCTGCAACGATTTGAGCGGGCCGACAAAGACTAGCGGCCGGAACCGCACCTCATCCAG 2700
Db 2660 AATCTGCAACGATTTGAGCGGGCCGACAAAGACTAGCGGCCGGAACCGCACCTCATCCAG 2719
QY 2701 GGCACCTGGACTTCTGTGATGGAATCTCTGTGTTCATCTCTCCGCGAGATGGCTTC 2760
Db 2720 GGCACCTGGACTTCTGTGATGGAATCTCTGTGTTCATCTCTCCGCGAGATGGCTTC 2779
QY 2761 AATGACTGTATGCAACAAAGGCGAGTGTGGGAGCTGTGCTTGCCTATCCCGGGCGC 2820
Db 2780 AATGACTGTATGCAACAAAGGCGAGTGTGGGAGCTGTGCTTGCCTATCCCGGGCGC 2839
QY 2821 CACCGCTGGGCTCGGCTCACACTACACCTTGACCCGACGAGCGCAACTGACGACCG 2880
Db 2840 CACCGCTGGGCTCGGCTCACACTACACCTTGACCCGACGAGCGCAACTGACGACCG 2899
QY 2881 CCCACCACTTCTTGTCTTTCAGCCAGAAATCTGCCATCTAGTCGAGATGATCCCGACGAC 2940
Db 2900 CCCACCACTTCTTGTCTTTCAGCCAGAAATCTGCCATCTAGTCGAGATGATCCCGACGAC 2959
QY 2941 CAGCACAGCCGGATCTCATCTCTGCGCTGATGAGCTGAGGAGCTCAAGACCATCGAC 3000
Db 2960 CAGCACAGCCGGATCTCATCTCTGCGCTGATGAGCTGAGGAGCTCAAGACCATCGAC 3019
QY 3001 TATGACCCACTGGACAAAGTTTCATCTACTGGGTGATGGGCCCGAGAAATCAAGCGAGCC 3060
Db 3020 TATGACCCACTGGACAAAGTTTCATCTACTGGGTGATGGGCCCGAGAAATCAAGCGAGCC 3079
QY 3061 AAGSACGAGGAGCCAGCGCTTGTGTGACCTCTCTGAGCCAAAGGCAAAACCCAGAC 3120
Db 3080 AAGSACGAGGAGCCAGCGCTTGTGTGACCTCTCTGAGCCAAAGGCAAAACCCAGAC 3139
QY 3121 AGGACGCGCCACGACCTCAGCATCGATCTTACAGCGGACACTGTTCTGGACGTCGAG 3180
Db 3140 AGGACGCGCCACGACCTCAGCATCGATCTTACAGCGGACACTGTTCTGGACGTCGAG 3199
QY 3181 GCCACCAATACCACTCAACGCTCCAGCGTGAAGGGGAGGCAATGGGGGTGGTCTGCTG 3240
Db 3200 GCCACCAATACCACTCAACGCTCCAGCGTGAAGGGGAGGCAATGGGGGTGGTCTGCTG 3259
QY 3241 GGGGACCGGACCAAGCCCGAGCCATCGTCTGTAACCGCGGAGCGAGGATCTGTGACTTC 3300
Db 3260 GGGGACCGGACCAAGCCCGAGCCATCGTCTGTAACCGCGGAGCGAGGATCTGTGACTTC 3319
QY 3301 ACCAAATGAGGACCGGGACGCAAGATCGAAACGCGGAGCCCTGGACGCGACCGAGCGC 3360
Db 3320 ACCAAATGAGGACCGGGACGCAAGATCGAAACGCGGAGCCCTGGACGCGACCGAGCGC 3379
QY 3361 GAGTCTCTTTCACCAACCGGCTCATCGCCCTGTGGCCCTGTGGTGGACAAACACTG 3420
Db 3380 GAGTCTCTTTCACCAACCGGCTCATCGCCCTGTGGCCCTGTGGTGGACAAACACTG 3439
QY 3421 GGCAGGCTGTCTGGGTGGACGCGGACCTGAAGCGCATTTGAGAGCTGTGACTGTGAGG 3480
Db 3440 GGCAGGCTGTCTGGGTGGACGCGGACCTGAAGCGCATTTGAGAGCTGTGACTGTGAGG 3499

QY	2281	GTGGCGGCTGGAGCGGCAGTTCCGGCAAGTCTCGTGTGGAGGACCTTGGACAAACCCG	2340
Db	2300	GTGGCGGCTGGAGCGGCAGTTCCGGCAAGTCTCGTGTGGAGGACCTTGGACAAACCCG	2359
QY	2341	AGGTGCTGGCGCTGGATCCCAACCAAGGGCTACATCTACTGGACCGAGTGGGGGGCAAG	2400
Db	2360	AGGTGCTGGCGCTGGATCCCAACCAAGGGCTACATCTACTGGACCGAGTGGGGGGCAAG	2419
QY	2401	CCGAGGATGTGCGGGCTTTCATGGA CCGGACCAACTCATGATGACGCTGGTGGACAAGGTG	2460
Db	2420	CCGAGGATGTGCGGGCTTTCATGGA CCGGACCAACTCATGATGACGCTGGTGGACAAGGTG	2479
QY	2461	GGCGGGCCACAGCACTCACCATTGACTACCTGACAGCGCCCTACTGACCGACCTG	2520
Db	2480	GGCGGGCCACAGCACTCACCATTGACTACCTGACAGCGCCCTACTGACCGACCTG	2539
QY	2521	GACACCAACATGATCGAGTCTGCAACATGCTGGGTGAGGAGCGGGTCTGATGCGCGAC	2580
Db	2540	GACACCAACATGATCGAGTCTGCAACATGCTGGGTGAGGAGCGGGTCTGATGCGCGAC	2599
QY	2581	GATCTCCCGCAACCGTTCCGCTGACGAGTACAGCGATTATATCTACTGGACAGACTGG	2640
Db	2600	GATCTCCCGCAACCGTTCCGCTGACGAGTACAGCGATTATATCTACTGGACAGACTGG	2659
QY	2641	AATCTGCACAGCATTTGAGCGGCGGACAAAGACTAGCGCGGGAACCGCACCTCATCCAG	2700
Db	2660	AATCTGCACAGCATTTGAGCGGCGGACAAAGACTAGCGCGGGAACCGCACCTCATCCAG	2719
QY	2701	GGCCACCTGGACTTCTGATGGAACATCTCTGTGTTCACACTCTCCCGCAGGATGGCTC	2760
Db	2720	GGCCACCTGGACTTCTGATGGAACATCTCTGTGTTCACACTCTCCCGCAGGATGGCTC	2779
QY	2761	AATGACTGATGCACAAACAGGGCAGTGTGGGAGCTGTGCTTGTGCCATCCCGCGGCG	2820
Db	2780	AATGACTGATGCACAAACAGGGCAGTGTGGGAGCTGTGCTTGTGCCATCCCGCGGCG	2839
QY	2821	CACCGCTGGGCTGGCTCACACTACACCTTGGACCCCGACGCGCAACTGACGCGCG	2880
Db	2840	CACCGCTGGGCTGGGCTCACACTACACCTTGGACCCCGACGCGCAACTGACGCGCG	2899
QY	2881	CCCACACTTCTTGTCTTTCAGCCAGAAATCTGCCATCAGTCTGGATGATCCCGAGCAG	2940
Db	2900	CCCACACTTCTTGTCTTTCAGCCAGAAATCTGCCATCAGTCTGGATGATCCCGAGCAG	2959
QY	2941	CAGCACGCGCGATCTCATCTGCGCTTGCATGGA CTGAGGAAGCTCAAGGCCATCGAC	3000
Db	2960	CAGCACGCGCGATCTCATCTGCGCTTGCATGGA CTGAGGAAGCTCAAGGCCATCGAC	3019
QY	3001	TATGACCACTGGACAAAGTTTCATCTGCTGGTGGATGGCGCCAGAACTCAAGCGGCC	3060
Db	3020	TATGACCACTGGACAAAGTTTCATCTGCTGGTGGATGGCGCCAGAACTCAAGCGGCC	3079
QY	3061	AAGGACGACGGGACCCAGCCCTTTTGTGACCTCTCTGAGCCAGGCGCAAAACCCAGAC	3120
Db	3080	AAGGACGACGGGACCCAGCCCTTTTGTGACCTCTCTGAGCCAGGCGCAAAACCCAGAC	3139
QY	3121	AGGACGCCGACGAGCTCAGCATCGACATCTACAGCCGGA CACTGTTCTGGACGTGCGAG	3180
Db	3140	AGGACGCCGACGAGCTCAGCATCGACATCTACAGCCGGA CACTGTTCTGGACGTGCGAG	3199
QY	3181	GCCACCAATACATCAACGTTCCACAGGCTGAGCGGGGAAGCCATGGGGGTGTGCTGGT	3240
Db	3200	GCCACCAATACATCAACGTTCCACAGGCTGAGCGGGGAAGCCATGGGGGTGTGCTGGT	3259
QY	3241	GGGACCGCGCAAGGCCAGGGCCATCGTCTGACCGGAGCGAGGGTACTGTACTTC	3300
Db	3260	GGGACCGCGCAAGGCCAGGGCCATCGTCTGACCGGAGCGAGGGTACTGTACTTC	3319
QY	3301	ACCAACATGACGAGCCGGGACCGAAGATCGAA CCGCGAGCGCTTGGACCGGACCGCGC	3360
Db	3320	ACCAACATGACGAGCCGGGACCGAAGATCGAA CCGCGAGCGCTTGGACCGGACCGCGC	3379

QY	3361	GAGTCTCTTTACACACCGGCTCATCGCGCTGTGTCGCCCTGTGTGTGGACAAACACACTG	3420
Db	3380	GAGTCTCTTTACACACCGGCTCATCGCGCTGTGTCGCCCTGTGTGTGGACAAACACACTG	3439
QY	3421	GGCAAGCTGTTCTGGGTGGACGCGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTGAGG	3480
Db	3440	GGCAAGCTGTTCTGGGTGGACGCGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTGAGG	3499
QY	3481	GCCAAACGCGCTGACCTTGGAGGACGCCAAACATCTGTGAGCCCTCTGGGCTGACCATCTT	3540
Db	3500	GCCAAACGCGCTGACCTTGGAGGACGCCAAACATCTGTGAGCCCTCTGGGCTGACCATCTT	3559
QY	3541	GGCAAGCATCTTACTGATCGACCGCGCAGCAGATGATCGAGCGTGTGAGAGAGACC	3600
Db	3560	GGCAAGCATCTTACTGATCGACCGCGCAGCAGATGATCGAGCGTGTGAGAGAGACC	3619
QY	3601	ACCGGGGACAAGCGGACTCGGCATCCAGGGCGGTGTGCGCCACCTCACTGGCATCCATGCA	3660
Db	3620	ACCGGGGACAAGCGGACTCGGCATCCAGGGCGGTGTGCGCCACCTCACTGGCATCCATGCA	3679
QY	3661	GTGAGGAGTTCAGCTTGGAGGAGTTCTCAGCCACCCATGTGCGCGTGA CAATGGTGGC	3720
Db	3680	GTGAGGAGTTCAGCTTGGAGGAGTTCTCAGCCACCCATGTGCGCGTGA CAATGGTGGC	3739
QY	3721	TGCTCCCACTCTCTGATTTGCGAAGGTGATGGGACACACAGGTGTCTATGCCAGTCCAC	3780
Db	3740	TGCTCCCACTCTCTGATTTGCGAAGGTGATGGGACACACAGGTGTCTATGCCAGTCCAC	3799
QY	3781	CTCTGTCTCTGCGAAGACCTGTGACCTGTGAGAGCGCGCCACCTGTCTCCCGGACCAAG	3840
Db	3800	CTCTGTCTCTGCGAAGACCTGTGACCTGTGAGAGCGCGCCACCTGTCTCCCGGACCAAG	3859
QY	3841	TTTTCATGTGCGCACAGGGGAGATCGACTGTATCCCGGGGCTCGGCGCTGTGAGCGCTT	3900
Db	3860	TTTTCATGTGCGCACAGGGGAGATCGACTGTATCCCGGGGCTCGGCGCTGTGAGCGCTT	3919
QY	3901	CCCAGTGCAGATGACAGAGCGAGGAGGCTGCGCGGTGTGTCTCGCGCGCCAGTTC	3960
Db	3920	CCCAGTGCAGATGACAGAGCGAGGAGGCTGCGCGGTGTGTCTCGCGCGCCAGTTC	3979
QY	3961	CCCTGCGCGGGGTCTAGTGTGTGGA CTTGCGCTGTGCGTGTGCGAGCGGAGGAGTGT	4020
Db	3980	CCCTGCGCGGGGTCTAGTGTGTGGA CTTGCGCTGTGCGTGTGCGAGCGGAGGAGTGT	4039
QY	4021	CAGACCGCTCAGACGAGGCGGACTGTGACGCCATCTGCGTCCCAACAGTTCGGGT	4080
Db	4040	CAGACCGCTCAGACGAGGCGGACTGTGACGCCATCTGCGTCCCAACAGTTCGGGT	4099
QY	4081	GCGAGCGGCCAGTGTCTCTCATCAAAACAGCAGTGTGCGACTCTTCCCGGACTGTATCGAC	4140
Db	4100	GCGAGCGGCCAGTGTCTCTCATCAAAACAGCAGTGTGCGACTCTTCCCGGACTGTATCGAC	4159
QY	4141	GGTCCCGACGAGCTCATGTGTGAATCA CCAAGCGCGCTTCTGAGAGCAGCGCGGCCAC	4200
Db	4160	GGTCCCGACGAGCTCATGTGTGAATCA CCAAGCGCGCTTCTGAGAGCAGCGCGGCCAC	4219
QY	4201	AGCAGTGCCTATCGGGCGGCTCATTTGGCATCATCTCTCTCTCTCTCTCTCTCTCTCTCT	4260
Db	4220	AGCAGTGCCTATCGGGCGGCTCATTTGGCATCATCTCTCTCTCTCTCTCTCTCTCTCTCT	4279
QY	4261	TATTTTGTGTGCCACGCGTGTGCGAGCGCTATGCGGGGGCAACCGGCGCTTCCG	4320
Db	4280	TATTTTGTGTGCCACGCGTGTGCGAGCGCTATGCGGGGGCAACCGGCGCTTCCG	4339
QY	4321	CACAGTATGTACGGGACCCCGCATCGTGCCTCTCAATTTATAGCCCGGGGGGTTC	4380
Db	4340	CACAGTATGTACGGGACCCCGCATCGTGCCTCTCAATTTATAGCCCGGGGGGTTC	4399
QY	4381	CAGCATGSCCTTTCACAGGCATCGCATGCGGAAAGTCCATGATGAGTCTCGTGAAGCTG	4440
Db	4400	CAGCATGSCCTTTCACAGGCATCGCATGCGGAAAGTCCATGATGAGTCTCGTGAAGCTG	4459
QY	4441	ATGGGGGGCGGGGGGGGGTGTCCCTCTGTACGACCGGAACCAAGTCCAGAGGGCGCTTCGTC	4500

QY 1081 CTGCTGCTGGCCCGCGGACGACCTACGAGGATCTCGTGGACACGCGGACTTACCC 1140
Db 1100 CTGCTGCTGGCCCGCGGACGACCTACGAGGATCTCGTGGACACGCGGACTTACCC 1159
QY 1141 GACATCGTGTGACGCTGACGACACATCCGCGACGCAATGCCATCGACTACGACCCGCTA 1200
Db 1160 GACATCGTGTGACGCTGACGACACATCCGCGACGCAATGCCATCGACTACGACCCGCTA 1219
QY 1201 GAGGCTATGCTTACTGGACAGATGACGAGTGTGGGCGCATCCGACGCGGTACTCTGGAC 1260
Db 1220 GAGGCTATGCTTACTGGACAGATGACGAGTGTGGGCGCATCCGACGCGGTACTCTGGAC 1279
QY 1261 GGGTCTGGGGCGCAGACCTGTGTAACACCGAGATCAAGACCCCGATGGCATCGCGTTC 1320
Db 1280 GGGTCTGGGGCGCAGACCTGTGTAACACCGAGATCAAGACCCCGATGGCATCGCGTTC 1339
QY 1321 GACTGGGTGGCCCGGAAACCTTACTGGACCGACACGCGACCGCATCGAGGTGACG 1380
Db 1340 GACTGGGTGGCCCGGAAACCTTACTGGACCGACACGCGACCGCATCGAGGTGACG 1399
QY 1381 GGCCTCAACGGGACCTCCGCAAGATCTGTGTGCGAGGACCTTGGACGAGCCCGAGCC 1440
Db 1400 GGCCTCAACGGGACCTCCGCAAGATCTGTGTGCGAGGACCTTGGACGAGCCCGAGCC 1459
QY 1441 ATCGCACTGCAACCCCGTGCATGGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAA 1500
Db 1460 ATCGCACTGCAACCCCGTGCATGGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAA 1519
QY 1501 ATCGAGTGTGCAACTTGGATGGGCGAGGCGGCTGTGCTCAATGCTCCCTCGGG 1560
Db 1520 ATCGAGTGTGCAACTTGGATGGGCGAGGCGGCTGTGCTCAATGCTCCCTCGGG 1579
QY 1561 TGGCCCAACGGGCTGGCCCTGACCTGACGAGGAGGAAAGTCTTACTTGGGAGAGCCCAAG 1620
Db 1580 TGGCCCAACGGGCTGGCCCTGACCTGACGAGGAGGAAAGTCTTACTTGGGAGAGCCCAAG 1639
QY 1621 ACAGACAGATCGAGTGCATCAATGTTGATGGAGCGAGAGGCGGACCTCTGGAGGAC 1680
Db 1640 ACAGACAGATCGAGTGCATCAATGTTGATGGAGCGAGAGGCGGACCTCTGGAGGAC 1699
QY 1681 AAGCTCCCGCACTTTTCGGGTTTCAAGCTGTGGGGGACTTTCATCTACTGCACTGACTGG 1740
Db 1700 AAGCTCCCGCACTTTTCGGGTTTCAAGCTGTGGGGGACTTTCATCTACTGCACTGACTGG 1759
QY 1741 CAGCGCCGACGATCGAGCGGGTGCACAAAGTCAAGGCCAGCGGGGACGTCAATATTGAC 1800
Db 1760 CAGCGCCGACGATCGAGCGGGTGCACAAAGTCAAGGCCAGCGGGGACGTCAATATTGAC 1819
QY 1801 CAGCTGCCCGACCTGATGGGGCTCAAAGCTGTGAATGTGGCCAAAGTCTGTGGACCAAC 1860
Db 1820 CAGCTGCCCGACCTGATGGGGCTCAAAGCTGTGAATGTGGCCAAAGTCTGTGGACCAAC 1879
QY 1861 CCGTGTGGGACAGGAAACGGGGGGTGCAGCACCTGTGCTTTCACACCCCAAGCAACC 1920
Db 1880 CCGTGTGGGACAGGAAACGGGGGGTGCAGCACCTGTGCTTTCACACCCCAAGCAACC 1939
QY 1921 CGGTGTGGCTGCCCATCGGCTGTGAGTGTGCTGAGTGAATGAAGACCTGTGCTGCT 1980
Db 1940 CGGTGTGGCTGCCCATCGGCTGTGAGTGTGCTGAGTGAATGAAGACCTGTGCTGCT 1999
QY 1981 GAGGCTTTCTTGTCTTCCACGACGAGCGCGCATCCACAGGATCTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTTTCTTGTCTTCCACGAGAGCGCGCATCCACAGGATCTCCCTCGAGACCAAT 2059
QY 2041 AACAAACGAGTGGCCATCCCGCTCAACGGGCGTCAAGGAGGCTCAGCCCTGGACTTTGAT 2100
Db 2060 AACAAACGAGTGGCCATCCCGCTCAACGGGCGTCAAGGAGGCTCAGCCCTGGACTTTGAT 2119
QY 2101 GTGTCCAAACCAACATCTACTGACAGAGCTGACGCTGAAGACCATCAGCGCGGCTTC 2160
Db 2120 GTGTCCAAACCAACATCTACTGACAGAGCTGACGCTGAAGACCATCAGCGCGGCTTC 2179

QY 2161 ATGAACGGGAGCTCGTGGAGCAGTGTGTGAGTTTGCCCTTGACTACCCCGAGGGCATG 2220
Db 2180 ATGAACGGGAGCTCGTGGAGCAGTGTGTGAGTTTGCCCTTGACTACCCCGAGGGCATG 2239
QY 2221 GCCGTGTGACTGGATGGGCAAGAACTCTTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2280
Db 2240 GCCGTGTGACTGGATGGGCAAGAACTCTTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2299
QY 2281 GTGGCGGCTGTGACGCGGACGTTCCGGCAAGTCTCTGTGTGAGGGGACTTGGACAAACCG 2340
Db 2300 GTGGCGGCTGTGACGCGGACGTTCCGGCAAGTCTCTGTGTGAGGGGACTTGGACAAACCG 2359
QY 2341 AGGTGCTGTGGCCCTGGATCCCAACAGGCTTACTACTGACCGGAGTGGGGGGCAG 2400
Db 2360 AGGTGCTGTGGCCCTGGATCCCAACAGGCTTACTACTGACCGGAGTGGGGGGCAG 2419
QY 2401 CCGAGGATCTGTGGCGGCTTTCATGACCGGACCAACTCTGATGACGCTGGTGACAAAGGTG 2460
Db 2420 CCGAGGATCTGTGGCGGCTTTCATGACCGGACCAACTCTGATGACGCTGGTGACAAAGGTG 2479
QY 2461 GGCGGGGCAACGACCTTCAACATTTGAATTAACGCTGACAGCGCTCTACTGACCGACCTG 2520
Db 2480 GGCGGGGCAACGACCTTCAACATTTGAATTAACGCTGACAGCGCTCTACTGACCGACCTG 2539
QY 2521 GACACCAACATGATGAGTGTGCTGCTGACGAGTACAGGATTAATATCTACTTGGACAGACTGG 2580
Db 2540 GACACCAACATGATGAGTGTGCTGCTGACGAGTACAGGATTAATATCTACTTGGACAGACTGG 2599
QY 2581 GATCTCCCGCAACCCGTTTGGTCTGACGAGTACAGGATTAATATCTACTTGGACAGACTGG 2640
Db 2600 GATCTCCCGCAACCCGTTTGGTCTGACGAGTACAGGATTAATATCTACTTGGACAGACTGG 2659
QY 2641 AATCTGACAGCATGTAGCGGCGGACGACAAAGTACAGGCGGCGGACCGACCTCTATCCAG 2700
Db 2660 AATCTGACAGCATGTAGCGGCGGACGACAAAGTACAGGCGGCGGACCGACCTCTATCCAG 2719
QY 2701 GGCCACCTTGGACTTGTGATGAGACATCTCTGTGTTCACCTCTCCCGCAGGATGGCTC 2760
Db 2720 GGCCACCTTGGACTTGTGATGAGACATCTCTGTGTTCACCTCTCCCGCAGGATGGCTC 2779
QY 2761 AATGACTGTATGCAACAAACGGGCGAGTGTGGGAGCTGTGCTTGGCCATCCCGGCGGC 2820
Db 2780 AATGACTGTATGCAACAAACGGGCGAGTGTGGGAGCTGTGCTTGGCCATCCCGGCGGC 2839
QY 2821 CACGCTGCGGCTGGCTTACACTACACCTTGAACCCGACGCGGACCGCACTGACGCGC 2880
Db 2840 CACGCTGCGGCTGGCTTACACTACACCTTGAACCCGACGCGGACCGCACTGACGCGC 2899
QY 2881 CCCACCACTTGTGTGCTTGTGAGCAGAAATCTGCCATCAGTCCGATGATCCCGGACGAC 2940
Db 2900 CCCACCACTTGTGTGCTTGTGAGCAGAAATCTGCCATCAGTCCGATGATCCCGGACGAC 2959
QY 2941 CAGCACAGCCGAGTCTCATCTCTGCCCTTGCATGGACTGAGGAAACGTCAAAGCCATCGAC 3000
Db 2960 CAGCACAGCCGAGTCTCATCTCTGCCCTTGCATGGACTGAGGAAACGTCAAAGCCATCGAC 3019
QY 3001 TATGACCCACTGGAACAAGTTTCATCTACTGGTGTGATGGGCGGACGACATCAAGGAGCC 3060
Db 3020 TATGACCCACTGGAACAAGTTTCATCTACTGGTGTGATGGGCGGACGACATCAAGGAGCC 3079
QY 3061 AAGGACGAGCGGACCCGAGCCCTTTTGTGACTTCTGTAGCCAAAGGCGGACCAACCCAGAC 3120
Db 3080 AAGGACGAGCGGACCCGAGCCCTTTTGTGACTTCTGTAGCCAAAGGCGGACCAACCCAGAC 3139
QY 3121 AGGACGCCCCACGACTTACGATCGACATCTTACAGCGGACACTGTGTTTGTGACGTTGCGAG 3180
Db 3140 AGGACGCCCCACGACTTACGATCGACATCTTACAGCGGACACTGTGTTTGTGACGTTGCGAG 3199
QY 3181 GCCACCAATACCAATCAAGTCCACAGGCTGAGCGGAGGAGCCATGGGGGTGTTGCTGCT 3240
Db 3200 GCCACCAATACCAATCAAGTCCACAGGCTGAGCGGAGGAGCCATGGGGGTGTTGCTGCT 3259
QY 3241 GGGGACCGCGACAAAGCCAGGGCCATCTGCTCAACGCGGAGCGAGGGGTACTCTGTACTTC 3300

Db	3260	GGGGA	CGCCGAC	AAAGCC	CAGGGCC	ATCGT	CGT	CAACG	CGAGGAG	GGT	TACCT	TGT	GTA	CTTC	3319														
Qy	3301	ACCAA	CATG	CAGGAC	CGGGG	CAGC	CAAGAT	CGAA	CGCG	CAGC	CCCT	TGGAC	GCAC	CCG	CAGCGC	3360													
Db	3320	ACCAA	CATG	CAGGAC	CGGGG	CAGC	CAAGAT	CGAA	CGCG	CAGC	CCCT	TGGAC	GCAC	CCG	CAGCGC	3379													
Qy	3361	GAGG	TCTCT	TTAC	ACAC	CGGCG	CTCAT	CCG	CCCT	TGT	TGGCC	CTGT	GTG	FTG	FGA	CAAC	CATG	3420											
Db	3380	GAGG	TCTCT	TTAC	CAAC	CGGCG	CTCAT	CCG	CCCT	TGT	TGGCC	CTGT	GTG	FTG	FGA	CAAC	CATG	3439											
Qy	3421	GGCAA	CGCT	TTCT	GGGT	GGA	CGGGA	CCCT	CAAG	CGCAT	TTG	CAG	AGCT	TG	GAC	CTGT	CAGGG	3480											
Db	3440	GGCAA	CGCT	TTCT	GGGT	GGA	CGGGA	CCCT	CAAG	CGCAT	TTG	CAG	AGCT	TG	GAC	CTGT	CAGGG	3499											
Qy	3481	GCCAA	CGCCT	GAC	CCCT	TGGAG	AGCG	CCAA	CAT	CGT	TG	CAG	CGCT	CT	TGG	GCCT	TG	CAAT	CTCT	3540									
Db	3500	GCCAA	CGCCT	GAC	CCCT	TGGAG	AGCG	CCAA	CAT	CGT	TG	CAG	CGCT	CT	TGG	GCCT	TG	CAAT	CTCT	3559									
Qy	3541	GGCAA	GCA	TC	TTA	CTG	GAT	TCGA	CCG	CAG	CAG	CAG	ATG	ATC	GAG	CGT	TG	FGA	AGAAC	3600									
Db	3560	GGCAA	GCA	TC	TTA	CTG	GAT	TCGA	CCG	CAG	CAG	CAG	ATG	ATC	GAG	CGT	TG	FGA	AGAAC	3619									
Qy	3601	ACCG	GGG	ACA	AG	CGG	ACT	TCG	CA	TCC	CAG	GGG	CCGT	TG	CCCA	CCCT	CAT	CGCC	ATCC	ATGCA	3660								
Db	3620	ACCG	GGG	ACA	AG	CGG	ACT	TCG	CA	TCC	CAG	GGG	CCGT	TG	CCCA	CCCT	CAT	CGCC	ATCC	ATGCA	3679								
Qy	3661	GTG	GAG	GA	CT	CAG	CC	TG	GAG	GA	TTCT	CAG	CC	CA	CC	CA	TG	CCG	TG	ACA	ATG	TGGC	3720						
Db	3680	GTG	GAG	GA	CT	CAG	CC	TG	GAG	GA	TTCT	CAG	CC	CA	CC	CA	TG	CCG	TG	ACA	ATG	TGGC	3739						
Qy	3721	TGCT	CC	CA	TC	TG	TAT	TG	CAAG	GGT	GTAT	GGA	CA	CA	CG	GTG	CT	CAT	GCC	CAG	TCC	AC	3780						
Db	3740	TGCT	CC	CA	TC	TG	TAT	TG	CAAG	GGT	GTAT	GGA	CA	CA	CG	GTG	CT	CAT	GCC	CAG	TCC	AC	3799						
Qy	3781	CTCG	TG	CT	CTG	CAG	AA	CT	TG	TGA	CT	TG	FGA	AG	CG	CCCA	CT	TG	CT	CCCC	CG	ACC	AG	3840					
Db	3800	CTCG	TG	CT	CTG	CAG	AA	CT	TG	TGA	CT	TG	FGA	AG	CG	CCCA	CT	TG	CT	CCCC	CG	ACC	AG	3859					
Qy	3841	TTTG	CA	TG	TGCC	ACA	G	GGG	AG	AT	CGA	CT	GT	AT	CCC	CG	GG	CC	TG	GG	CG	CT	GT	GA	CG	CT	TT	3900	
Db	3860	TTTG	CA	TG	TGCC	ACA	G	GGG	AG	AT	CGA	CT	GT	AT	CCC	CG	GG	CC	TG	GG	CG	CT	GT	GA	CG	CT	TT	3919	
Qy	3901	CCG	AG	TG	CG	AT	GAC	AC	GAG	CG	AG	GGG	GT	CG	CC	CG	TG	CT	CG	CG	CG	CC	CAG	TT	3960				
Db	3920	CCG	AG	TG	CG	AT	GAC	AC	GAG	CG	AG	GGG	GT	CG	CC	CG	TG	CT	CG	CG	CG	CC	CAG	TT	3979				
Qy	3961	CCCT	TG	CG	CG	GGG	T	CA	GT	GT	G	B	A	C	T	TG	G	B	A	C	T	TG	G	B	A	C	T	GT	4020
Db	3980	CCCT	TG	CG	CG	GGG	T	CA	GT	GT	G	B	A	C	T	TG	G	B	A	C	T	TG	G	B	A	C	T	GT	4039
Qy	4021	CAG	AC	CG	CT	C	A	G	AC	G	G	CG	GA	CT	GT	G	A	C	G	CG	CA	CT	GT	G	A	C	CG	GT	4080
Db	4040	CAG	AC	CG	CT																								

RESULTS 5

RESULTS 3
IIS-09-949-016-2457

US-09-3439-016-2437 : Sequence 2457. Application IIS/09949016

Sequence 2437, Appl.
: Patent No 6812339

; PATENT NO. 6812339
: GENERAL INFORMATION

; GENERAL INFORMATION:
: APPLICANT: VENTER J Craig et al

APPLICANT: VENTER, J. Craig et al.

1. TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE METHODS OF DETECTION AND USES THEREOF

; TITLE OF INVENTION: WITH

FILE REFERENCE: CL001307
CURRENT IDENTIFICATION NUMBER: US/00/040 016

; CURRENT APPLICATION NUMBER: US/0

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

;
; PRIOR FILING DATE: 2000-10-20

;
PRIOR APPLICATION NUMBER: 60/

; PRIOR FILING DATE: 2000-10-03

60/

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: Fast

; SEQ ID NO 2457

; LENGTH: 5

; ORGANISM: Human
US-09-949-016-2457

Query Match 99.6%; Score 5042.8; DB 4; Length 5103;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 5058; Conservative 0; Mismatches 2; Indels 4;

[illegible]

QY 2104 TCCAAACACCATCTACTGACAGACGTCAGCCTGGAAGACCATCAGCGCGCCTTCATG 2163
Db 2101 TCCAAACACCATCTACTGACAGACGTCAGCCTGGAAGACCATCAGCGCGCCTTCATG 2160
QY 2164 AACGGGAGCTCGGTGGAGCAAGTGTGAGTTTGGCTTGTACTACCCAGAGGCAATGGCC 2223
Db 2161 AACGGGAGCTCGGTGGAGCAAGTGTGAGTTTGGCTTGTACTACCCAGAGGCAATGGCC 2220
QY 2224 GTTGACTGGATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTG 2283
Db 2221 GTTGACTGGATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTG 2280
QY 2284 GCGGGCTTGGACGGGCAAGTTCGGCAAGTCTCGTGTGGAGGGACTTGGACCAACCCGAGG 2343
Db 2281 GCGGGCTTGGACGGGCAAGTTCGGCAAGTCTCGTGTGGAGGGACTTGGACCAACCCGAGG 2340
QY 2344 TCGCTGGCCCTGGATCCCAACCAAGGGCTTACATCTACTGGACCGAGTGGGGCGGCAAGCCG 2403
Db 2341 TCGCTGGCCCTGGATCCCAACCAAGGGCTTACATCTACTGGACCGAGTGGGGCGGCAAGCCG 2400
QY 2404 AGGATCGTGGCGGCTTCCATGGACGGGACCAACTGCATGACGCTGGTGGAACAAGGTGGGC 2463
Db 2401 AGGATCGTGGCGGCTTCCATGGACGGGACCAACTGCATGACGCTGGTGGAACAAGGTGGGC 2460
QY 2464 CCGGCCAACGACCTCACCATTTGACTACGCTGACAGCGCCTCTACTGGACCGACCTGGAC 2523
Db 2461 CCGGCCAACGACCTCACCATTTGACTACGCTGACAGCGCCTCTACTGGACCGACCTGGAC 2520
QY 2524 ACCAATGATGATGATGTCCTCAACATGCTGGGTGAGGAGCGGGTGTGATTTGCCGACGAT 2583
Db 2521 ACCAATGATGATGATGTCCTCAACATGCTGGGTGAGGAGCGGGTGTGATTTGCCGACGAT 2580
QY 2584 CTCGGGACCGCTTCGGTCTCAGCCAGTACAGCGATATATATCTACTGACAGACTGGAT 2643
Db 2581 CTCGGGACCGCTTCGGTCTCAGCCAGTACAGCGATATATATCTACTGACAGACTGGAT 2640
QY 2644 CTGCACAGCATTTGAGCGGGCGGCAAGACTAGCGGCGGAAACCGCACCTCTATCCAGGGC 2703
Db 2641 CTGCACAGCATTTGAGCGGGCGGCAAGACTAGCGGCGGAAACCGCACCTCTATCCAGGGC 2700
QY 2704 CACTGGAATTTGTGATGGAATCTGTGTGTTCATCTCTCCCGCAGGATGGCTCAAT 2763
Db 2701 CACTGGAATTTGTGATGGAATCTGTGTGTTCATCTCTCCCGCAGGATGGCTCAAT 2760
QY 2764 GACTGTATGCAACAACAGGGCAGTGTGGCAGCTGTGTGCTTGCATCTCCCGGCGGCAC 2823
Db 2761 GACTGTATGCAACAACAGGGCAGTGTGGCAGCTGTGTGCTTGCATCTCCCGGCGGCAC 2820
QY 2824 CGCTGCGGCTCGGCTCACACTACACCTTGACGCCAGCAGCGCAACTGCGAGCCGCGCC 2883
Db 2821 CGCTGCGGCTCGGCTCACACTACACCTTGACGCCAGCAGCGCAACTGCGAGCCGCGCC 2880
QY 2884 ACCACCTTTTGTGTTGTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGGACGACGAG 2943
Db 2881 ACCACCTTTTGTGTTGTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGGACGACGAG 2940
QY 2944 CACAGCCGGATCTCATCTCGCCCTGCAATGCACTGAGGACGTCGAAGCCATCGACTAT 3003
Db 2941 CACAGCCGGATCTCATCTCGCCCTGCAATGCACTGAGGACGTCGAAGCCATCGACTAT 3000
QY 3004 GACCCACTGGACAAGTTTCATCTACTGGGTGGATGGGCGCCAGAAACATCAAGCGAGCCAAAG 3063
Db 3001 GACCCACTGGACAAGTTTCATCTACTGGGTGGATGGGCGCCAGAAACATCAAGCGAGCCAAAG 3060
QY 3064 GACGACGGGACCCAGGGCAGCCCTTGTGTTTGGACCTCTCTGAGCCAAAGCCAAACCCAGA 3119
Db 3061 GACGACGGGACCCAGGGCAGCCCTTGTGTTTGGACCTCTCTGAGCCAAAGCCAAACCCAGA 3120
QY 3120 CAGGACGGCCCAACGACCTCAGCATGCACTTACAGCCGGACACTGTTCTGGAGCTGGGA 3179
Db 3121 CAGGACGGCCCAACGACCTCAGCATGCACTTACAGCCGGACACTGTTCTGGAGCTGGGA 3180

QY 3180 GGCCACCAATACCATCAACGCTCCACAGGCTGAGGGGGAAGCCATGGGGGTGGTCTGCG 3239
Db 3181 GGCCACCAATACCATCAACGCTCCACAGGCTGAGCGGGGAAGCCATGGGGGTGGTCTGCG 3240
QY 3240 TGGGGAACGCGCAAGGCCAGGCGCATGCTGCTCAACGCGGAGGAGGGTACTCTGTACTT 3299
Db 3241 TGGGGAACGCGCAAGGCCAGGCGCATGCTGCTCAACGCGGAGGAGGGTACTCTGTACTT 3300
QY 3300 CACCAACATGACGAGACCGGCGAGCCAGATCGAACGCGCAGCCCTGGACGGACCGAGCG 3359
Db 3301 CACCAACATGACGAGACCGGCGAGCCAGATCGAACGCGCAGCCCTGGACGGACCGAGCG 3360
QY 3360 CGAGGCTCTCTTACCAACCGGCTCATCCGCGCTGTGGCCCTGTGGTGGCAACACACT 3419
Db 3361 CGAGGCTCTCTTACCAACCGGCTCATCCGCGCTGTGGCCCTGTGGTGGCAACACACT 3420
QY 3420 GGGCAAGCTGTTCTGGGTGGACGGGACCTGAAGCGCATTTGAGAGCTGTGACTCTGACG 3479
Db 3421 GGGCAAGCTGTTCTGGGTGGACGGGACCTGAAGCGCATTTGAGAGCTGTGACTCTGACG 3480
QY 3480 GGCCAAACGCGCTGACCCCTGGAGGACGCCAACATCGTCAGCCTCTGGGCTTGACCATCT 3539
Db 3481 GGCCAAACGCGCTGACCCCTGGAGGACGCCAACATCGTCAGCCTCTGGGCTTGACCATCT 3540
QY 3540 TGGCAAGCATCTTACTGGATCGACCGCAGCAGCAGATGATCGAGCCTGTGGGAAGAC 3599
Db 3541 TGGCAAGCATCTTACTGGATCGACCGCAGCAGCAGATGATCGAGCCTGTGGGAAGAC 3600
QY 3600 CACCGGGGAACAGCGGACTTGCATCCAGGGCGGTGTGCCCACTCTACTGGGATCCATGCG 3659
Db 3601 CACCGGGGAACAGCGGACTTGCATCCAGGGCGGTGTGCCCACTCTACTGGGATCCATGCG 3660
QY 3660 AGTCGAGGAGTCAAGCTGGAGGAGTTCTCAGCCACCCATGTCGCCCTGACCAATGGTGG 3719
Db 3661 AGTCGAGGAGTCAAGCTGGAGGAGTTCTCAGCCACCCATGTCGCCCTGACCAATGGTGG 3720
QY 3720 CTGCTCCACATCTGTATTTGCCAAGGGTGTGGGACACACCGGTGCTCATGCCAGTCCA 3779
Db 3721 CTGCTCCACATCTGTATTTGCCAAGGGTGTGGGACACACCGGTGCTCATGCCAGTCCA 3780
QY 3780 CCTCGTGTCTTGACAACTGTGACTGTGGAGAGCGGCCACCTGCTCTCCCGGACCA 3839
Db 3781 CCTCGTGTCTTGACAACTGTGACTGTGGAGAGCGGCCACCTGCTCTCCCGGACCA 3840
QY 3840 GTTTGCAATGTCACACGAGGAGATCGACTGTATCCCGGGGCTGGGCGCTGTGACGGCTT 3899
Db 3841 GTTTGCAATGTCACACGAGGAGATCGACTGTATCCCGGGGCTGGGCGCTGTGACGGCTT 3900
QY 3900 TCCCGAGTGGCATGACACGAGCGACGAGGAGGCTGCGCCCTGTGCTCCGCGGCCAGTT 3959
Db 3901 TCCCGAGTGGCATGACACGAGCGACGAGGAGGCTGCGCCCTGTGCTCCGCGGCCAGTT 3960
QY 3960 CCCCTGCGCGGGGTCAAGTGTGGACCTTGGCGCTGTGGCGAGCGGAGGAGAGCTG 4019
Db 3961 CCCCTGCGCGGGGTCAAGTGTGGACCTTGGCGCTGTGGCGAGCGGAGGAGAGCTG 4020
QY 4020 TCAGGACCGCTCAGACGAGCGGAGCTGTGAGCCCATCTGCGCTGCCCAACAGTTCGGGTG 4079
Db 4021 TCAGGACCGCTCAGACGAGCGGAGCTGTGAGCCCATCTGCGCTGCCCAACAGTTCGGGTG 4080
QY 4080 TGCAGCGGCGCAGTGTGCTCTCATCAACAGCAGTGGAGCTCTTCCCGGAGCTGTATCGA 4139
Db 4081 TGCAGCGGCGCAGTGTGCTCTCATCAACAGCAGTGGAGCTCTTCCCGGAGCTGTATCGA 4140
QY 4140 CGGCTCCGACGAGTCAATGTGTGAAATCAACGAGCGCCCTCAGACGAGCCCGGCCCA 4199
Db 4141 CGGCTCCGACGAGTCAATGTGTGAAATCAACGAGCGCCCTCAGACGAGCCCGGCCCA 4200
QY 4200 CAGCAGTGGCATCGGGCCGCTCATTTGGCATCATCTCTCTCTCTGTCATGGGTGGTGT 4259
Db 4201 CAGCAGTGGCATCGGGCCGCTCATTTGGCATCATCTCTCTCTCTCTGTCATGGGTGGTGT 4260
QY 4260 CTATTTTGTGTGCCAGCGCGTGGTGTGCCAGCGCTATGCGGGGCGCAACGCGGCCCTTCCC 4319

	Db	4261	CTATTTTGTGTGCAGGCGGTGGTGTGCACGGCTATGTGCGGGGCCAAACGGGGCCCTTCCC	4320
	Qy	4320	GCACGAGTATGTACAGCGGGAACCCCGCACGTGCCCTCAAATTTATAGCCCGCGCGGTTC	4379
	Db	4321	GCACGAGTATGTACAGCGGACCCCGACGTGCCCTCAAATTTATAGCCCGCGCGGTTC	4380
	Qy	4380	CCAGCATGGCCCTTTCACAGCATCGCATGCGGAAAGTCCATGATAGCTCCCGTAGACCT	4439
	Db	4381	CCAGCATGGCCCTTTCACAGCATCGCATGCGGAAAGTCCATGATAGCTCCCGTAGACCT	4440
	Qy	4440	GATGGGGGCGGGGCGGGGTGCCCTGTAGCACCGGAACCAAGTCAACAGGGGCTCTGTC	4499
	Db	4441	GATGGGGGCGGGGCGGGGTGCCCTGTAGCACCGGAACCAAGTCAACAGGGGCTCTGTC	4500
	Qy	4500	CAGCAGCTCTGTCAGACGCAAGGCCACGCTGTACCCGCCCATCTCTGAACCCGCGGCCCTC	4559
	Db	4501	CAGCAGCTCTGTCAGACGCAAGGCCACGCTGTACCCGCCCATCTCTGAACCCGCGGCCCTC	4560
	Qy	4560	CCCGGCCACGGACCCCTCCCTGTACAACTAGGACATGTTCTACTCTTCAAACATTCGGC	4619
	Db	4561	CCCGGCCACGGACCCCTCCCTGTACAACTAGGACATGTTCTACTCTTCAAACATTCGGC	4620
	Qy	4620	CATCTGCAGACCGTACAGGCCCTACATCATTCGAGGAATGGCGCCCGCACGCGCCCTG	4679
	Db	4621	CATCTGCAGACCGTACAGGCCCTACATCATTCGAGGAATGGCGCCCGCACGCGCCCTG	4680
	Qy	4680	CAGACCCGACGTGTGTACAGCAGCTACAGGCCCGCCGCTGGAAGCCAGCAAGTACTA	4739
	Db	4681	CAGACCCGACGTGTGTACAGCAGCTACAGGCCCGCCGCTGGAAGCCAGCAAGTACTA	4740
	Qy	4740	CCTGGATTTGAACTCGGACTCAGACCCCTATCCACCCCGCACCGCCCGCACAGCCAGTA	4799
	Db	4741	CCTGGATTTGAACTCGGACTCAGACCCCTATCCACCCCGCACCGCCCGCACAGCCAGTA	4800
	Qy	4800	CCTGTGCGCGGAGACAGTCTCCCGCCCTCGCCCGCCACCGAGAGGAGTACTTCCATCT	4859
	Db	4801	CCTGTGCGCGGAGACAGTCTCCCGCCCTCGCCCGCCACCGAGAGGAGTACTTCCATCT	4860
	Qy	4860	CTTCCCGCCCCCTCCGTCGCCCTGTGACGAGTCACTCTGACCTCGGCGGGCCACTCTGG	4919
	Db	4861	CTTCCCGCCCCCTCCGTCGCCCTGTGACGAGTCACTCTGACCTCGGCGGGCCACTCTGG	4920
	Qy	4920	CTTCTCTGTGCCCTGTAAATAGTTTAAATATGAACAAGAAAAATATATTTTATGA	4979
	Db	4921	CTTCTCTGTGCCCTGTAAATAGTTTAAATATGAACAAGAAAAATATATTTTATGA	4980
	Qy	4980	TTTTAAAAATAAATAAATTTGGGATTTAAAAACATGAGAAATGTCAACTGTGATGGGT	5039
	Db	4981	TTTTAAAAATAAATAAATTTGGGATTTAAAAACATGAGAAATGTCAACTGTGATGGGT	5040
	Qy	5040	GGGCGGGCTGGGAGAACTTTGTA 5063	
	Db	5041	GGGCGGGCTGGGAGAACTTTGTA 5064	

RESULT 6

```

US-09-060-299-1
; Sequence 1, Application US/09060299
; Patent No. 6545137
; GENERAL INFORMATION:
;
; APPLICANT: Todd, John A
; APPLICANT: Hess, John W
; APPLICANT: Caskey, Charles T
; APPLICANT: Cox, Roger D
; APPLICANT: Gerhold, David
; APPLICANT: Hammond, Holly
; APPLICANT: Hey, Patricia
; APPLICANT: Kawaguchi, Yoshihiko
; APPLICANT: Merzhan, Tony R
; APPLICANT: Metzker, Michael L
; TITLE OF INVENTION: No. 6545137el Receptor
; NUMBER OF SEQUENCES: 455

```

Db 481 |||||CTCTTCTGCGAGGACCTTTGACCGAGCGGAGCCATCGCCTTGGACCCCGCTCAGCGGTAC 540
QY 544 ATGTACTGGA CAGACTGGGTGGAGACGCCCGCGATTGAGCGGGCAGGATGATGGCAGC 603
Db 541 ATGTACTGGA CAGACTGGGTGGAGACGCCCGCGATTGAGCGGGCAGGATGATGGCAGC 600
QY 604 ACCCGGAAGATCATTTGTGGA CTGGA CATTTACTGGCCCAATGGACTGACCATGACACTG 663
Db 601 ACCCGGAAGATCATTTGTGGA CTGGA CATTTACTGGCCCAATGGACTGACCATGACACTG 660
QY 664 GAGGAGCAGAAGCTTACTTGGCTGACGCCAAGCTCAGCTTCATCCACCGTGCACACTG 723
Db 661 GAGGAGCAGAAGCTTACTTGGCTGACGCCAAGCTCAGCTTCATCCACCGTGCACACTG 720
QY 724 GACGCTCGTTCCGCGCAGAGGTGTGGAGGGCAGCTGACGCCACCCCTTGCCTCTGAG 783
Db 721 GACGCTCGTTCCGCGCAGAGGTGTGGAGGGCAGCTGACGCCACCCCTTGCCTCTGAG 780
QY 784 CTCTCCGGGACACTCTCTACTGGA CAGACTGGCAGACCCCGCTCCATCCATGCTGCAAC 843
Db 781 CTCTCCGGGACACTCTCTACTGGA CAGACTGGCAGACCCCGCTCCATCCATGCTGCAAC 840
QY 844 AAGCGCACTGGGGGGAAGAGGAGATCTGAGTGCCCTCTACTCACCATGAGACATC 903
Db 841 AAGCGCACTGGGGGGAAGAGGAGATCTGAGTGCCCTCTACTCACCATGAGACATC 900
QY 904 CAGGTGCTGACCGCAGGAGCGCGCTTTCTTCCACACTGCTGTGAGGAGACATGGC 963
Db 901 CAGGTGCTGACCGCAGGAGCGCGCTTTCTTCCACACTGCTGTGAGGAGACATGGC 960
QY 964 GECTGCTCCCACTGTGCTGTCTGCCCAAGCAGAGCTTTTACACATGCGCTGCCCC 1023
Db 961 GECTGCTCCCACTGTGCTGTCTGCCCAAGCAGAGCTTTTACACATGCGCTGCCCC 1020
QY 1024 ACGGGTGTGAGCTGTGAGCA CAA CGGCAGGACGTGTAAAGCAGGAGCGGAGGTGCTG 1083
Db 1021 ACGGGTGTGAGCTGTGAGCA CAA CGGCAGGACGTGTAAAGCAGGAGCGGAGGTGCTG 1080
QY 1084 CTGCTGGCCCGGCGAGCGACTTACGGAGGATCTCGCTGGACACCGCGGACTTACCGAC 1143
Db 1081 CTGCTGGCCCGGCGAGCGACTTACGGAGGATCTCGCTGGACACCGCGGACTTACCGAC 1140
QY 1144 ATCGTGTGAGGTGAGCAGCATCGGCACGCCATTGGCATCGACTAGACCCGCTAGAG 1203
Db 1141 ATCGTGTGAGGTGAGCAGCATCGGCACGCCATTGGCATCGACTAGACCCGCTAGAG 1200
QY 1204 GGCTATGTCTACTGGACAGATGACGAGGTGCGGGCCATCCGACAGGCGGTACTTGGACGG 1263
Db 1201 GGCTATGTCTACTGGACAGATGACGAGGTGCGGGCCATCCGACAGGCGGTACTTGGACGG 1260
QY 1264 TCTGGGGCGCAGCGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGGTGAC 1323
Db 1261 TCTGGGGCGCAGCGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGGTGAC 1320
QY 1324 TGGGTGGCCCAAACTCTACTGACACGACAGCGGSCAGCGACCGCATCGAGGTGACGGCG 1383
Db 1321 TGGGTGGCCCAAACTCTACTGACACGACAGCGGSCAGCGACCGCATCGAGGTGACGGCG 1380
QY 1384 CTC AACCGCACCTCCCGCAAGATCCTGTGTGTCGAGGACCTTGGACGAGCCCGCGAGCCATC 1443
Db 1381 CTC AACCGCACCTCCCGCAAGATCCTGTGTGTCGAGGACCTTGGACGAGCCCGCGAGCCATC 1440
QY 1444 GCATGCA CCCCCTGATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAAATC 1503
Db 1441 GCATGCA CCCCCTGATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAAATC 1500
QY 1504 GAGTGTGCAACTTGGATGGGCGAGCGGCTGTGCTGGTCAATGCTCCCTCGGTTGG 1563
Db 1501 GAGTGTGCAACTTGGATGGGCGAGGAGCGGCTGTGCTGGTCAATGCTCCCTCGGTTGG 1560
QY 1564 CCCAACCGCTTGGCCCTGGACCTG CAGGAGGGGAGCTCTACTGGGAGAGCCCAAGACA 1623

Db 1561 CCCAACCGCTTGGCCCTGGACCTG CAGGAGGGGAAGCTCTACTGGGGAGAGCCCAAGACA 1620
QY 1624 GACAAGATCGAGGTGATCAATGTTGATGGGACGAGAGCGGACCTCTCTGGAGACAAAG 1683
Db 1621 GACAAGATCGAGGTGATCAATGTTGATGGGACGAGAGCGGACCTCTCTGGAGACAAAG 1680
QY 1684 CTC CGGCACATTTTTCGGGTTCA CGCTGCTGGGGACTTCACTACTGAGTACTGACTGGCAG 1743
Db 1681 CTC CGGCACATTTTTCGGGTTCA CGCTGCTGGGGACTTCACTACTGAGTACTGACTGGCAG 1740
QY 1744 CGCGCAGCATCGAGCGGGTGCA CAAAGGTCAAGCCAGCCGAGCGGACGTCACTATTGACAG 1803
Db 1741 CGCGCAGCATCGAGCGGGTGCA CAAAGGTCAAGCCAGCCGAGCGGACGTCACTATTGACAG 1800
QY 1804 CTGCGCCGACCTGATGGGGCTCAAAGCTGTGAAATGTGGCCAAAGGTGCTGCGAACCAACCG 1863
Db 1801 CTGCGCCGACCTGATGGGGCTCAAAGCTGTGAAATGTGGCCAAAGGTGCTGCGAACCAACCG 1860
QY 1864 TGTGCGGACAGGAACGGGGGGTGCA GCCACCTGTGCTCTTTCACACCCCAAGCAACCCGG 1923
Db 1861 TGTGCGGACAGGAACGGGGGGTGCA GCCACCTGTGCTCTTTCACACCCCAAGCAACCCGG 1920
QY 1924 TGTGGCTGCCCATCGGCTGGAGCTGCTGAGTGACATGAAGACTGTGATCTGCTGAG 1983
Db 1921 TGTGGCTGCCCATCGGCTGGAGCTGCTGAGTGACATGAAGACTGTGATCTGCTGAG 1980
QY 1984 GCCTTCTTGGTCTTTCACAGCAGAGCGGCCATCCACAGGATCTCCCTCGAGACCAATAAC 2043
Db 1981 GCCTTCTTGGTCTTTCACAGCAGAGCGGCCATCCACAGGATCTCCCTCGAGACCAATAAC 2040
QY 2044 AACGACGTGGCCATCCCGCTCAGCGGGCTCAAGGAGGCTCAGGCCCTGGACTTTGATGTG 2103
Db 2041 AACGACGT -GCCATCCCGCTCAGCGGGCTCAAGGAGGCTCAGGCCCTGGACTTTGATGTG 2099
QY 2104 TCCAAACCAACATCTACTGGA CAGAGCTGAGCTGAAAGCAATCAGCGCGCTTCATG 2163
Db 2100 TCCAAACCAACATCTACTGGA CAGAGCTGAGCTGAAAGCAATCAGCGCGCTTCATG 2159
QY 2164 AACGGGAGCTCGGTGGAGCAGCTGCTGGAGTTTGGCCTTGACTACCCGAGGGCATGCC 2223
Db 2160 AACGGGAGCTCGGTGGAGCAGCTGCTGGAGTTTGGCCTTGACTACCCGAGGGCATGCC 2219
QY 2224 GTTGACTGGATGGGCAAGAACCTCTACTGGGCCGACACTGGGACCAACAGAAATCGAAATG 2283
Db 2220 GTTGACTGGATGGGCAAGAACCTCTACTGGGCCGACACTGGGACCAACAGAAATCGAAATG 2279
QY 2284 GCGCGCTGGA CCGGCAAGTTCCGGCAAGTCTCTGTGTGGAGGGA CTTGGACAACCCGAGG 2343
Db 2280 GCGCGCTGGA CCGGCAAGTTCCGGCAAGTCTCTGTGTGGAGGGA CTTGGACAACCCGAGG 2339
QY 2344 TCGCTGGCCCTGGATCCCA CCAAGGGCTTACATCTACTGGACCGGATGGGGCGGCAAGCGG 2403
Db 2340 TCGCTGGCCCTGGATCCCA CCAAGGGCTTACATCTACTGGACCGGATGGGGCGGCAAGCGG 2399
QY 2404 AGGATCGTGGGGCTTTCATGGA CCGGACCAACCTGCTGCTGCTGGTGGGACCAAGGTGGGC 2463
Db 2400 AGGATCGTGGGGCTTTCATGGA CCGGACCAACCTGCTGCTGCTGGTGGGACCAAGGTGGGC 2459
QY 2464 CGGGCCAAACCACTTCA CCAATTTGACTTACGCTGGA CCGGCTCTTACTGGACCGACTGAC 2523
Db 2460 CGGGCCAAACCACTTCA CCAATTTGACTTACGCTGGA CCGGCTCTTACTGGACCGACTGAC 2519
QY 2524 ACCAAATGATCGAGTCTGTC CCAATGCTGGGTGAGGAGCGGGTGGTATTGCCGACGAT 2583
Db 2520 ACCAAATGATCGAGTCTGTC CCAATGCTGGGTGAGGAGCGGGTGGTATTGCCGACGAT 2579
QY 2584 CTC CGGCACCCGTTTCGGTCTGGA CCGAGTACAGCGATTATATCTACTGGACAGACTGGAAT 2643
Db 2580 CTC CGGCACCCGTTTCGGTCTGGA CCGAGTACAGCGATTATATCTACTGGACAGACTGGAAT 2639
QY 2644 CTGCA CAGCATTTGAGCGGGCGCA CAAAGACTAGCGGCGGGAACCCGACCTCATCCAGGGC 2703
Db 2640 CTGCA CAGCATTTGAGCGGGCGCA CAAAGACTAGCGGCGGGAACCCGACCTCATCCAGGGC 2699

QY 2704 CACCTGGACTTCGTGATGAGACATCTCTGGTGTTCACACTCTCTCCGCGCAGGATGGCCTCAAT 2763
Db 2700 CACCTGGACTTCGTGATGAGACATCTCTGGTGTTCACACTCTCTCCGCGCAGGATGGCCTCAAT 2759
QY 2764 GACTGTATGACAAACAAACGGGCGAGTGTGGGAGCTGTGCTTGCCTCCGCGCGGCCAC 2823
Db 2760 GACTGTATGACAAACAAACGGGCGAGTGTGGGAGCTGTGCTTGCCTCCGCGCGGCCAC 2819
QY 2824 CGCTGGGTGGCCTTCACTACACCTTGACCCAGCAGCCGCAATGTCAGCCCGCC 2883
Db 2820 CGCTGGGTGGCCTTCACTACACCTTGACCCAGCAGCCGCAATGTCAGCCCGCC 2879
QY 2884 ACCACTCTTCTGCTGTGAGCCAGAAATCTGCCATCAGTCGGATGATCCGAGACACAG 2943
Db 2880 ACCACTCTTCTGCTGTGAGCCAGAAATCTGCCATCAGTCGGATGATCCGAGACACAG 2939
QY 2944 CACAGCCGGATCTCATCTCTGCCCTGCGATGAGTGTGAGGAACTGTCAGGACCTCAACTAT 3003
Db 2940 CACAGCCGGATCTCATCTCTGCCCTGCGATGAGTGTGAGGAACTGTCAGGACCTCAACTAT 2999
QY 3004 GACCACTGGACAAATTCATCTACTTGGGTGAGTGGCGCGAGAAATCAAGCGACCAAG 3063
Db 3000 GACCACTGGACAAATTCATCTACTTGGGTGAGTGGCGCGAGAAATCAAGCGACCAAG 3059
QY 3064 GACGACGGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAGGCGCAAAACCCAGACAG 3123
Db 3060 GACGACGGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAGGCGCAAAACCCAGACAG 3119
QY 3124 CAGCCCCACGACCTCAGATCGACATCTACAGCCGAGACACTGTTCTGAGCGTGCAGGCC 3183
Db 3120 CAGCCCCACGACCTCAGATCGACATCTACAGCCGAGACACTGTTCTGAGCGTGCAGGCC 3179
QY 3184 ACCAATACATCAAGCTTCAAGGCTGAGCGGGGAAGCCATGCGGGTGGTGTGGTGG 3243
Db 3180 ACCAATACATCAAGCTTCAAGGCTGAGCGGGGAAGCCATGCGGGTGGTGTGGTGG 3239
QY 3244 GACCGCGAACAAGCCAGGGCCATCTGCTCAAGCGGAGCGAGGTACTCTGACTTACC 3303
Db 3240 GACCGCGAACAAGCCAGGGCCATCTGCTCAAGCGGAGCGAGGTACTCTGACTTACC 3299
QY 3304 AACATGCGAGGACCGGGCAGCGAAGTCAAGCGCAGCCCTGGAAGCGGACCGAGCGCAG 3363
Db 3300 AACATGCGAGGACCGGGCAGCGAAGTCAAGCGCAGCCCTGGAAGCGGACCGAGCGCAG 3359
QY 3364 GTCTCTTTCACACCGGCTCATCCGCGCTGTGGCCCTGTGGTGGAGCAACACACTGGGC 3423
Db 3360 GTCTCTTTCACACCGGCTCATCCGCGCTGTGGCCCTGTGGTGGAGCAACACACTGGGC 3419
QY 3424 AAGCTGTTCTGGGTGGACCGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTCAAGGGGCC 3483
Db 3420 AAGCTGTTCTGGGTGGACCGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTCAAGGGGCC 3479
QY 3484 AACCGCTGACCTGGAGGACCGCAACATCTGTGAGCCTCTGGGCTGACCATCTTTGGC 3543
Db 3480 AACCGCTGACCTGGAGGACCGCAACATCTGTGAGCCTCTGGGCTGACCATCTTTGGC 3539
QY 3544 AAGCATCTTACTGATCGACCGCGCAGCAGCATGATGATGAGCTGTGGAGAAACACACC 3603
Db 3540 AAGCATCTTACTGATCGACCGCGCAGCAGCATGATGATGAGCTGTGGAGAAACACACC 3599
QY 3604 GGGGACAGCGGACTCGCATCCAGGGCCGTGTGCGCCACCTCACTGGCATCCATGCAATG 3663
Db 3600 GGGGACAGCGGACTCGCATCCAGGGCCGTGTGCGCCACCTCACTGGCATCCATGCAATG 3659
QY 3664 GAGGAAGTCAGCTGGAGGATTTCTACGCCACCCATGTGCCGTGAGCAATGTGGTGGTC 3723
Db 3660 GAGGAAGTCAGCTGGAGGATTTCTACGCCACCCATGTGCCGTGAGCAATGTGGTGGTC 3719
QY 3724 TCCACATCTGATTTGCCAAGGATGAGGACACACAGGCTCATGCGCCAGTCCACCTC 3783
Db 3720 TCCACATCTGATTTGCCAAGGATGAGGACACACAGGCTCATGCGCCAGTCCACCTC 3779

QY 3784 GTGCTCTTGAGAACCTGCTGACCTGTGTGAGAGCGCGCCACCTGCTCCCGGACAGTTT 3843
Db 3780 GTGCTCTTGAGAACCTGCTGACCTGTGTGAGAGCGCGCCACCTGCTCCCGGACAGTTT 3839
QY 3844 GCATGTGACACAGGGAGATCGACTGTATCCCGGGGCTGGCGCTGTGAGCGCTTTCC 3903
Db 3840 GCATGTGACACAGGGAGATCGACTGTATCCCGGGGCTGGCGCTGTGAGCGCTTTCC 3899
QY 3904 GAGTGCATGACACAGAGCGACGAGAGGGCTGCCCGGTGTCTCCGCGCCAGTTCC 3963
Db 3900 GAGTGCATGACACAGAGCGACGAGAGGGCTGCCCGGTGTCTCCGCGCCAGTTCC 3959
QY 3964 TGCGCGGGGTGAGTGTGTGAGCTTGCCTGTGCTGCGAGCGGAGGACGTGTGAG 4023
Db 3960 TGCGCGGGGTGAGTGTGTGAGCTTGCCTGTGCTGCGAGCGGAGGACGTGTGAG 4019
QY 4024 GACCGCTCAGACGAGGGGAGCTGTGAGCGCCATCTGCTGCCCAACACAGTTCCGGTGTGCG 4083
Db 4020 GACCGCTCAGACGAGGGGAGCTGTGAGCGCCATCTGCTGCCCAACACAGTTCCGGTGTGCG 4079
QY 4084 AGCGGCGAGTGTGCTCATCAAAACAGAGTGTGAGCTCTTCCCGGACTGTATGACGCGC 4143
Db 4080 AGCGGCGAGTGTGCTCATCAAAACAGAGTGTGAGCTCTTCCCGGACTGTATGACGCGC 4138
QY 4144 TCCGACGAGCTCATGTGTGAAATCACCAAGCGCCCTCAGACGACAGCCCGGCCACAGC 4203
Db 4139 TCCGACGAGCTCATGTGTGAAATCACCAAGCGCCCTCAGACGACAGCCCGGCCACAGC 4198
QY 4204 AGTGCCATCGGCGCCGCTCATTTGGCATCATCTCTCTCTCTTCTGTCATGGGTGTCTAT 4263
Db 4199 AGTGCCATCGGCGCCGCTCATTTGGCATCATCTCTCTCTCTCTGTCATGGGTGTCTAT 4258
QY 4264 TTTGTGTGCGAGCGCGTGTGTGCGAGCTATGCGGGGGCAACGGGCGCTTCCCGCAC 4323
Db 4259 TTTGTGTGCGAGCGCGTGTGTGCGAGCTATGCGGGGGCAACGGGCGCTTCCCGCAC 4318
QY 4324 GAGTATGTGACGCGGACCCCGACGTGCTCAATTTCTAGAGCCCGGGGCTTCCCG 4383
Db 4319 GAGTATGTGACGCGGACCCCGACGTGCTCAATTTCTAGAGCCCGGGGCTTCCCG 4378
QY 4384 CATGGCCCTTTCACAGGCATCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCTGATG 4443
Db 4379 CATGGCCCTTTCACAGGCATCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCTGATG 4438
QY 4444 GGGGGCGGGGCGGGGTGCCCTGTACGACGGAACCACTGTCAGAGGGCTTCTGCGCAGC 4503
Db 4439 GGGGGCGGGGCGGGGTGCCCTGTACGACGGAACCACTGTCAGAGGGCTTCTGCGCAGC 4498
QY 4504 AGCTCGTCCAGCACAAAGGCGACGCTGTACCGCGCGATCTGAAACCGCGCCCTCCCGC 4563
Db 4499 AGCTCGTCCAGCACAAAGGCGACGCTGTACCGCGCGATCTGAAACCGCGCCCTCCCGC 4558
QY 4564 GCCACGACCCCTCTCTGTACAAACATGGAATGTTTACTTCTTCAAAATTCGCGGCACT 4623
Db 4559 GCCACGACCCCTCTCTGTACAAACATGGAATGTTTACTTCTTCAAAATTCGCGGCACT 4618
QY 4624 GCGAGACGCTACAGGCGCTCATCATCTTCGAGGAATGCGGCGCCCGGAGCGCTGCGCAGC 4683
Db 4619 GTGAGACGCTACAGGCGCTCATCATCTTCGAGGAATGCGGCGCCCGGAGCGCTGCGCAGC 4678
QY 4684 ACCGACGCTGTGACACGAGCTACAGCGCGACGCGCTGGAAGGCGCAGCAAGTACTACTG 4743
Db 4679 ACCGACGCTGTGACACGAGCTACAGCGCGCAGCGCTGGAAGGCGCAGCAAGTACTACTG 4738
QY 4744 GATTTGAATCGGACTCAGACCCCTATTCACCCCAACCCAGCGCCCAAGCAAGTACTG 4803
Db 4739 GATTTGAATCGGACTCAGACCCCTATTCACCCCAACCCAGCGCCCAAGCAAGTACTG 4798
QY 4804 TCGCGGAGGACAGCTGCGCGCCCTCGCGCCCAACCCAGAGGAGTACTTCCATCTCTTC 4863
Db 4799 TCGCGGAGGACAGCTGCGCGCCCTCGCGCCCAACCCAGAGGAGTACTTCCATCTCTTC 4858
QY 4864 CCGCCCCCTCCGCTCCGCTGACGAGCTCATCTGAGCTCGCGCGGGGCCACTCTGGGTTTC 4923

1084 QY CTGCTGGCCCGCGGACGACCTTACGAGGATCTCGCTGGACACCGCGGACTTTCACGAC 1143
1081 Db CTGCTGGCCCGCGGACGACCTTACGAGGATCTCGCTGGACACCGCGGACTTTCACGAC 1140
1144 QY ATCGTGTCTGAGTGGACGACATCCGGGACGACCTTGGCCATCGACTAGACCCGCTAGAG 1203
1141 Db ATCGTGTCTGAGTGGACGACATCCGGGACGACCTTGGCCATCGACTAGACCCGCTAGAG 1200
1204 QY GCGTATGTCTACTGACACGATGACGAGGTGGGCGCATCCCGAGCGGCTGACTGACACCGG 1263
1201 Db GCGTATGTCTACTGACACGATGACGAGGTGGGCGCATCCCGAGCGGCTGACTGACACCGG 1260
1264 QY TCTGGGGCGGACAGCCTGGTCAACACCGAGATCAACGACCCGATGGGATCGCGTGCAC 1323
1261 Db TCTGGGGCGGACAGCCTGGTCAACACCGAGATCAACGACCCGATGGGATCGCGTGCAC 1320
1324 QY TGGGTGGCCCGAAACCTTCTACTGACCGACACCGGACCGGACCGCATCGAGGTGACGCGC 1383
1321 Db TGGGTGGCCCGAAACCTTCTACTGACCGACACCGGACCGGACCGCATCGAGGTGACGCGC 1380
1384 QY CTCACCGGACCTCCCGCAAGATCTGTGTGCGAGGACCTGAGACGACCCCGAGCCATC 1443
1381 Db CTCACCGGACCTCCCGCAAGATCTGTGTGCGAGGACCTGAGACGACCCCGAGCCATC 1440
1444 QY GCACCTGACCCCGTGTGATGGGCTCATGTACTGACAGACTGGGAGAGAACCTTAATATC 1503
1441 Db GCACCTGACCCCGTGTGATGGGCTCATGTACTGACAGACTGGGAGAGAACCTTAATATC 1500
1504 QY GAGTGTGCAACTTGGATGGGACGAGCGGGCTGTGTGCTGCTCAATGCTCCCTCGGGTGG 1563
1501 Db GAGTGTGCAACTTGGATGGGACGAGCGGGCTGTGTGCTGCTCAATGCTCCCTCGGGTGG 1560
1564 QY CCCAACGCGCTGGCCCTGGAACCTGACGAGGAGGAAAGCTCTACTGAGGAGAGACCCCAAGACA 1623
1561 Db CCCAACGCGCTGGCCCTGGAACCTGACGAGGAGGAAAGCTCTACTGAGGAGAGACCCCAAGACA 1620
1624 QY GACAAAGATCGAGGTGATCAATGTTGATGGGACGAGGCGGACCTCTCTGGAGGACAAAG 1683
1621 Db GACAAAGATCGAGGTGATCAATGTTGATGGGACGAGGCGGACCTCTCTGGAGGACAAAG 1680
1684 QY CTCGCGCACATTTTCGGGTTTCAACGCTGTGGGGACTTTCATCTACTGAGTCACTGGGCGAG 1743
1681 Db CTCGCGCACATTTTCGGGTTTCAACGCTGTGGGGACTTTCATCTACTGAGTCACTGGGCGAG 1740
1744 QY CGCGCAGCATCGAGCGGGTGCACAAAGTCAAGGCTCAAGGCTCAAGGCTCAATTTGACCCAG 1803
1741 Db CGCGCAGCATCGAGCGGGTGCACAAAGTCAAGGCTCAAGGCTCAAGGCTCAATTTGACCCAG 1800
1804 QY CTGCGCGACCTGATGGGCTCAAAAGCTGTGAATGTGGCCAAAGTCTGCGGAACCAACCCG 1863
1801 Db CTGCGCGACCTGATGGGCTCAAAAGCTGTGAATGTGGCCAAAGTCTGCGGAACCAACCCG 1860
1864 QY TGTGCGGACAGGAAACGGGGGTGACGCCACCTGTGCTTCTTCAACACCCGACGCAACCCGG 1923
1861 Db TGTGCGGACAGGAAACGGGGGTGACGCCACCTGTGCTTCTTCAACACCCGACGCAACCCGG 1920
1924 QY TGTGCTGCCCATCGGCTGAGAGTGTGATGACATGAAAGACCTTGATCTGTCGCTGAG 1983
1921 Db TGTGCTGCCCATCGGCTGAGAGTGTGATGACATGAAAGACCTTGATCTGTCGCTGAG 1980
1984 QY GCCTTCTTGGTCTTCAACGACGAGCGCCCATCCACAGGATCTCCCTCGAGACCAATAAC 2043
1981 Db GCCTTCTTGGTCTTCAACGACGAGCGCCCATCCACAGGATCTCCCTCGAGACCAATAAC 2040
2044 QY AACGACGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 2103
2041 Db AACGACGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 2099
2104 QY TCCACACACCATCTACTGACAGAGCTGAGCTGAGCTGAGACCATCAGCGCGGCTTTCATG 2163
2100 Db TCCACACACCATCTACTGACAGAGCTGAGCTGAGCTGAGACCATCAGCGCGGCTTTCATG 2159

2164 QY AACCGGAGCTCGGTGGAGCAGCTGTGTGAGTTTGGCTTGTACTACCCCGAGGGCATGGCC 2223
2160 Db AACCGGAGCTCGGTGGAGCAGCTGTGTGAGTTTGGCTTGTACTACCCCGAGGGCATGGCC 2219
2224 QY GTTGACTGAGTGGGCAAGAACTTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAAGTG 2283
2220 Db GTTGACTGAGTGGGCAAGAACTTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAAGTG 2279
2284 QY GCGGCGTGGACGCGGAGTTCGCGCAAGTCTCTGTGTGGAGGGACTTGGACAAACCCGAGG 2343
2280 Db GCGGCGTGGACGCGGAGTTCGCGCAAGTCTCTGTGTGGAGGGACTTGGACAAACCCGAGG 2339
2344 QY TCGTGTGCGCTGGATCCCAACAGGCTTACATCTACTGGAACGAGTGGGCGGAGCGG 2403
2340 Db TCGTGTGCGCTGGATCCCAACAGGCTTACATCTACTGGAACGAGTGGGCGGAGCGG 2399
2404 QY AGGATCTGTGCGGCGCTTTCATGGAACGAGGCTTTCATGGAACGAGTGGGCGGAGCGG 2463
2400 Db AGGATCTGTGCGGCGCTTTCATGGAACGAGGCTTTCATGGAACGAGTGGGCGGAGCGG 2459
2464 QY GCGGCGCAACGACCTTTCATGGAACGAGGCTTTCATGGAACGAGTGGGCGGAGCGG 2523
2460 Db GCGGCGCAACGACCTTTCATGGAACGAGGCTTTCATGGAACGAGTGGGCGGAGCGG 2519
2524 QY ACCAACATGATCGAGTCTGCAACATGCTGGGTGAGGAGCGGCTGATTTGCCGACGAT 2583
2520 Db ACCAACATGATCGAGTCTGCAACATGCTGGGTGAGGAGCGGCTGATTTGCCGACGAT 2579
2584 QY CTCGCGCACCCGTTTGGTCTGACGAGTACAGGCAATATATCTACTGGAACGAGTGGAAAT 2643
2580 Db CTCGCGCACCCGTTTGGTCTGACGAGTACAGGCAATATATCTACTGGAACGAGTGGAAAT 2639
2644 QY CTGACAGCATTTGAGCGGGCGGACAAAGATAGCGGCGGAAACCGGACCTCATCAGGGC 2703
2640 Db CTGACAGCATTTGAGCGGGCGGACAAAGATAGCGGCGGAAACCGGACCTCATCAGGGC 2699
2704 QY CACCTGAGCTTCTGTGATGGACATCTGCTGTTTCACTCTCCCGCAGGATGGCTCAAT 2763
2700 Db CACCTGAGCTTCTGTGATGGACATCTGCTGTTTCACTCTCCCGCAGGATGGCTCAAT 2759
2764 QY GACTGTATGCAACAAACGCGGAGTGTGGGAGCTGTGCTTGGCATCTCCCGGCGGCGAC 2823
2760 Db GACTGTATGCAACAAACGCGGAGTGTGGGAGCTGTGCTTGGCATCTCCCGGCGGCGAC 2819
2824 QY CGCTGCGGCTGCGGCTCACACTACACCTGGAACCGGACCGGACCTGAGCGCGGCGG 2883
2820 Db CGCTGCGGCTGCGGCTCACACTACACCTGGAACCGGACCGGACCTGAGCGCGGCGG 2879
2884 QY ACCACCTTCTGTGTTTTCAGCAAGAAATCTGCAATCAGTGGATGATCCCGGACGACGAG 2943
2880 Db ACCACCTTCTGTGTTTTCAGCAAGAAATCTGCAATCAGTGGATGATCCCGGACGACGAG 2939
2944 QY CACAGCCGAGTCTCATCTCTGCGGCTGCACTGGAATGAGGAAACGTCAAAGCCATCGACTAT 3003
2940 Db CACAGCCGAGTCTCATCTCTGCGGCTGCACTGGAATGAGGAAACGTCAAAGCCATCGACTAT 2999
3004 QY GACCCACTGAGCAAGTTCATCTACTGAGTGGGCGGACGAGCAACATCAGCGAGGCGAG 3063
3000 Db GACCCACTGAGCAAGTTCATCTACTGAGTGGGCGGACGAGCAACATCAGCGAGGCGAG 3059
3064 QY GACGAGCGGACCCAGCGCTTGTGTTTTCAGCTCTCTGAGCGCAAGGCGGAGGCGGAGG 3123
3060 Db GACGAGCGGACCCAGCGCTTGTGTTTTCAGCTCTCTGAGCGCAAGGCGGAGGCGGAGG 3119
3124 QY CAGCCCGACGACCTCAGCATCGACATCTACAGCGGACACTGTTTCTGGAGCTGCGAGGCG 3183
3120 Db CAGCCCGACGACCTCAGCATCGACATCTACAGCGGACACTGTTTCTGGAGCTGCGAGGCG 3179
3184 QY ACCAATACCATCAACGCTGACGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 3243
3180 Db ACCAATACCATCAACGCTGACGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 3239
3244 QY GACCGGACCAAGCCCGAGGCGGCGATCGTCTGTCACACGCGGAGCGAGGCTACCTGTACTTACC 3303

Db 3240 GACCGCGAAGCCAGCGCCATCGTCTCAAGCGGAGCGAGGTACTGTGTTCCACC 3299
QY 3304 AACATGAGGACCGGGAGCCCAAGATCGAAGCGCGAGCCCTGGAGCGGACCGAGCGGAG 3363
Db 3300 AACATGAGGACCGGGAGCCCAAGATCGAAGCGCGAGCCCTGGAGCGGACCGAGCGGAG 3359
QY 3364 GTCTCTTTCAACACCGGCTCATCGCCCTGTGGCCCTGTGGTGGAGCAACACACTGGGC 3423
Db 3360 GTCTCTTTCAACACCGGCTCATCGCCCTGTGGCCCTGTGGTGGAGCAACACACTGGGC 3419
QY 3424 AAGCTGTTCTGGTGGAGCGCGAAGCTGAAGCGCAATTGAGAGCTGTGACCTGTCAAGGGCC 3483
Db 3420 AAGCTGTTCTGGTGGAGCGCGAAGCTGAAGCGCAATTGAGAGCTGTGACCTGTCAAGGGCC 3479
QY 3484 AACCGCTGACCTGGAGAGCGCAACATCGTGAGCCCTGTGGCCCTGACCATCTTGGC 3543
Db 3480 AACCGCTGACCTGGAGAGCGCAACATCGTGAGCCCTGTGGCCCTGACCATCTTGGC 3539
QY 3544 AAGCATCTTACTGATCGAAGCGCGAGCAGAGATGATCGAGCGGTGTGGAGAGACACACC 3603
Db 3540 AAGCATCTTACTGATCGAAGCGCGAGCAGAGATGATCGAGCGGTGTGGAGAGACACACC 3599
QY 3604 GGGACAAAGCGAAGCTCGCATCAAGGGCGGTGTGGCCCACTCACTGGCATCCATGCAAGT 3663
Db 3600 GGGACAAAGCGAAGCTCGCATCAAGGGCGGTGTGGCCCACTCACTGGCATCCATGCAAGT 3659
QY 3664 GAGGAAGTCAAGCGGAGGAGTCTCAGCCCAACCATGTGCGGCTGAGCAATGGTGGCTGC 3723
Db 3660 GAGGAAGTCAAGCGGAGGAGTCTCAGCCCAACCATGTGCGGCTGAGCAATGGTGGCTGC 3719
QY 3724 TCCACATCTGTATTGCAAGGAGTGTGGAGACCAAGCGGTCTCATGCCAGTCCACCTC 3783
Db 3720 TCCACATCTGTATTGCAAGGAGTGTGGAGACCAAGCGGTCTCATGCCAGTCCACCTC 3779
QY 3784 GTGCTCTCGAAGACCTGCTGACCTGTGGAGAGCGGCCACCTGCTCCCGGACCAAGTTT 3843
Db 3780 GTGCTCTCGAAGACCTGCTGACCTGTGGAGAGCGGCCACCTGCTCCCGGACCAAGTTT 3839
QY 3844 GCATGTGCCACAGGGAGATGATGATATCCCGGGGCTGGCGTGTGAGCGCTTCCC 3903
Db 3840 GCATGTGCCACAGGGAGATGATGATATCCCGGGGCTGGCGTGTGAGCGCTTCCC 3899
QY 3904 GAGTGCATGACAGAGCGAGGAGGCTGCCCGTGTGCTCCCGCGCCAGTCCC 3963
Db 3900 GAGTGCATGACAGAGCGAGGAGGCTGCCCGTGTGCTCCCGCGCCAGTCCC 3959
QY 3964 TGC CGCGGGGTCAAGTGTGGAAGCTGCGCTGCGCTGCGAGCGGAGGAGCTGTGAG 4023
Db 3960 TGC CGCGGGGTCAAGTGTGGAAGCTGCGCTGCGCTGCGAGCGGAGGAGCTGTGAG 4019
QY 4024 GACCGCTCAGAGGGGAGCTGTGAGCGCATCTGCTGCCCAACAGTTCGGTGTGCG 4083
Db 4020 GACCGCTCAGAGGGGAGCTGTGAGCGCATCTGCTGCCCAACAGTTCGGTGTGCG 4079
QY 4084 AGCGCGAGTGTGCTCATCAAGCAGAGTGTGAGTCTTCCCGAGCTGTGAGCGGC 4143
Db 4080 AGCGCGAGTGTGCTCATCAAGCAGAGTGTGAGTCTTCCCGAGCTGTGAGCGGC 4138
QY 4144 TCCGACGAGCTCATGTGTGAATCAACAGCGGCTTCAGACGAGCGCGGCCACAGC 4203
Db 4139 TCCGACGAGCTCATGTGTGAATCAACAGCGGCTTCAGACGAGCGCGGCCACAGC 4198
QY 4204 AGTGCATCGGGCGCTCATTTGGCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4263
Db 4199 AGTGCATCGGGCGCTCATTTGGCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4258
QY 4264 TTTGTGTCAGCGCGTGTGCGAGCGCTATGCGGGGCAAGCGGCCCTTCCCGCAC 4323
Db 4259 TTTGTGTCAGCGCGTGTGCGAGCGCTATGCGGGGCAAGCGGCCCTTCCCGCAC 4318
QY 4324 GAGTATGTGAGGGGACCCCGCAGCGTGCCTCTCAATTTATAGCCCGGGCGGTTCGCCAG 4383

Db 4319 GAGTATGTGAGGGGACCCCGCAGCGTGCCTCTCAATTTATAGATCCCGGGCGGTTCGCCAG 4378
QY 4384 CATGGCCCTTCACAGGATCGCATCGGAAAGTCCATGATGAGTCCGTGAGCTGATG 4443
Db 4379 CATGGCCCTTCACAGGATCGCATCGGAAAGTCCATGATGAGTCCGTGAGCTGATG 4438
QY 4444 GGGGGCGGGGCGGGGTGCCCTCTAGACCGGAAACACGTCACAGGGGCTTCGTCCAGC 4503
Db 4439 GGGGGCGGGGCGGGGTGCCCTCTAGACCGGAAACACGTCACAGGGGCTTCGTCCAGC 4498
QY 4504 AGTCTGTCCAGACGAGAGGCGACGCTGTACCCGCGGATTCCTGAAACCCCGCGCTTC 4563
Db 4499 AGTCTGTCCAGACGAGAGGCGACGCTGTACCCGCGGATTCCTGAAACCCCGCGCTTC 4558
QY 4564 GCCAGGACCCCTCCCTGTACCAATGAGACATGTTACTCTTCAAAATTCGCGGCACT 4623
Db 4559 GCCAGGACCCCTCCCTGTACCAATGAGACATGTTACTCTTCAAAATTCGCGGCACT 4618
QY 4624 GCGAGACCGTACAGGCGCTTACATCATTCGAGGAATGGCGCCCGGACGACGCGCTTCGAGC 4683
Db 4619 GTGAGACCGTACAGGCGCTTACATCATTCGAGGAATGGCGCCCGGACGACGCGCTTCGAGC 4678
QY 4684 ACCGAGTGTGTGACAGGAGCTACAGCGCGAGCGCTGGAAAGGCGAGCAAGTATACCTG 4743
Db 4679 ACCGAGTGTGTGACAGGAGCTACAGCGCGAGCGCTGGAAAGGCGAGCAAGTATACCTG 4738
QY 4744 GATTTGAATCGGACTCAGACCCCTATCCACCCCGGACGCGCCCGGACGAGCTACCTG 4803
Db 4739 GATTTGAATCGGACTCAGACCCCTATCCACCCCGGACGCGCCCGGACGAGCTACCTG 4798
QY 4804 TCGCGGAGGACAGCTCGCGCCCTCGCGCCCGGACGCGAGGAGCTACTTCCATCTCTTC 4863
Db 4799 TCGCGGAGGACAGCTCGCGCCCTCGCGCCCGGACGCGAGGAGCTACTTCCATCTCTTC 4858
QY 4864 CCGCCCTCGCTCGCTCGCGAGCTCATCTGAGCTCATCTGAGCTCGCGCGGCGCACTGCGCTTC 4923
Db 4859 CCGCCCTCGCTCGCTCGCGAGCTCATCTGAGCTCATCTGAGCTCGCGCGGCGCACTGCGCTTC 4918
QY 4924 TCTGTGCGCTGTAAATAGTTTAAATATGAACAAAGAAATATATATATATATATATAT 4983
Db 4919 TCTGTGCGCTGTAAATAGTTTAAATATGAACAAAGAAATATATATATATATATATAT 4978
QY 4984 AAAATAAATATATATGGGATTTTAAATAATGAGAAATGTGAATGTGATGGGTGGGC 5043
Db 4979 AAAATAAATATATATGGGATTTTAAATAATGAGAAATGTGAATGTGATGGGTGGGC 5038
QY 5044 AGGCTCGGAGAACTTTGTA 5063
Db 5039 AGGCTCGGAGAACTTTGTA 5058

RESULT 8

US-09-060-299-32
; Sequence 32, Application US/09060299
; Patent No. 6545137
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; APPLICANT: Hess, John W
; APPLICANT: Caskey, Charles T
; APPLICANT: Cox, Roger D
; APPLICANT: Gerhold, David
; APPLICANT: Hammond, Holly
; APPLICANT: Hey, Patricia
; APPLICANT: Kawaguchi, Yoshihiko
; APPLICANT: Merriman, Tony R
; APPLICANT: Metzker, Michael L
; TITLE OF INVENTION: No. 6545137el Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6545137th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia

COUNTRY: US
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/060,299
 FILING DATE: 15-APR-1998
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/043,553
 FILING DATE: 15-APR-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/048,740
 FILING DATE: 05-JUN-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: B.J.Sadoff
 REGISTRATION NUMBER: 36,663
 REFERENCE/DOCKET NUMBER: 620-35
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)816-4091
 TELEFAX: (703)816-4100
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5022 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-060-299-32

Query Match 96.8%; Score 4902; DB 4; Length 5022;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 4916; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY	142	GCCTCGCGCTCTCTGCTATTTCCAAACCGCGGACGCTACGCGCTGGTGGAGCCCGCGCGGA	201
DB	62	GCCTCGCGCTCTCTGCTATTTGCCAAACCGCGGACGCTACGCGTGGTGGAGCCCGCGCGGA	121
QY	202	GTCAGCTGGAGTACACATCGTGTGTCAGCGCCCTGGAGGATCGCGCGCGAGTGTGCTTC	261
DB	122	GTCAAGCTGGATCCACCATCGTGTGTCAGCGCCCTGGAGGATCGCGCGCGAGTGTGCTTC	181
QY	262	CAGTTTTCAGAGGAGCGCTGTACTGGACAGACGTGAGCGAGGAGCCATCAAGCAGACC	321
DB	192	CAGTTTTCNAGGAGCGGTGTACTGGACAGACG-GAGCGAGAGGCCATCAAGCAGACC	240
QY	322	TACCTGAACAGACGCGGCGCGCGTGTGAGAACTGTGCTATCTCGCGCTGTCTCTCCC	381
DB	241	TACCTGAACAGACGCGGCGCGCGTGTGAGAACTGTGCTATCTCGCGCTGTCTCTCCC	300
QY	382	GACGCGCTCGCTGCTGCTGGTGGGCAAGAGTGTACTGGACGAGCTCAGAGACCAAC	441
DB	301	GACGCGCTCGCTGCTGCTGGTGGGCAAGAGTGTACTGGACGAGCTCAGAGACCAAC	360
QY	442	CGCATCGAGGTGGCCAACTCAATGGCACATCCCGGAAGTGTCTCTTGGCAGGACCTT	501
DB	361	CGCATCGAGGTGGCCAACTCAATGGCACATCCCGGAAGTGTCTCTTGGCAGGACCTT	420
QY	502	GACGAGCTAGGGCCATCGCTTGGACCCCGCTCAGCGGTACATGTACTGGACAGACTGG	561
DB	421	GACGAGCTAGGGCCATCGCTTGGACCCCGCTCAGCGGTACATGTACTGGACAGACTGG	480
QY	562	GCTGAGACCGCCGATTTGAGCGGCGAGGATGGATGGACGACCCGGAAGATCATTTGTG	621
DB	481	GCTGAGACCGCCGATTTGAGCGGCGAGGATGGATGGACGACCCGGAAGATCATTTGTG	540
QY	622	GACTCGGACATTTACTGGCCCAATGGACTGACCATCGACCTGGAGGAGCAGAGCTCTAC	681
DB	541	GACTCGGACATTTACTGGCCCAATGGACTGACCATCGACCTGGAGGAGCAGAGCTCTAC	600
QY	682	TGGGCTGACGCGCTCAGCTTTCATCCACGCTGTCGCAACCTGGAGCGGCTGTTCCGGCAG	741

DB	601	TGGGCTGACGCGCAAGCTCAGCTTCATCCACGCTGCCAACCTGGAGCGGCTGTTCCGGCAG	660
QY	742	AAGGTGGTGGAGGCGCTTCAACGCAACCTTTCGCGCTTGACGCTCTCCGGGACACTCTG	801
DB	661	AAGGTGGTGGAGGCGCTTCAACGCAACCTTTCGCGCTTGACGCTCTCCGGGACACTCTG	720
QY	802	TACTGGACAGACTGGCAGACCGCTTCCATCATGCTGCAACAGCGCACTGGGGGGAAG	861
DB	721	TACTGGACAGACTGGCAGACCGCTTCCATCATGCTGCAACAGCGCACTGGGGGGAAG	780
QY	862	AGGAAGAGATCCTGAGTGCCTTCTACTCACCCATGGACATCCAGGTCTGAGCCAGAG	921
DB	781	AGGAAGAGATCCTGAGTGCCTTCTACTCACCCATGGACATCCAGGTCTGAGCCAGAG	840
QY	922	CGGACAGCCTTTTCCACATCTGCTGTGAGAGGACAATGGCGGTGTCTCCACCTGTGC	981
DB	841	CGGACAGCCTTTTCCACATCTGCTGTGAGAGGACAATGGCGGTGTCTCCACCTGTGC	900
QY	982	CTGCTGTCCCAAGGAGCCTTTCTACATGCGGCTTCCCAAGGCTGTCAGTGCAG	1041
DB	901	CTGCTGTCCCAAGGAGCCTTTCTACATGCGGCTTCCCAAGGCTGTCAGTGCAG	960
QY	1042	GACAAACGCGAGCGTGTAAAGCGAGGAGCGAGAGGTGTCTGCTGCGCCCGCGGAGC	1101
DB	961	GACAAACGCGAGCGTGTAAAGCGAGGAGCGAGAGGTGTCTGCTGCGCCCGCGGAGC	1020
QY	1102	GACCTAGCGAGGATCTGCTGGACACGCGCGGACTTCCACCGACATCTGCTGAGGTGAC	1161
DB	1021	GACCTAGCGAGGATCTGCTGGACACGCGCGGACTTTCACCGACATCTGCTGAGGTGAC	1080
QY	1162	GACATCCGCGACGCAATTTGCCATCGACTACGACCGCTTAGAGGGCTATGTCTACTGGAC	1221
DB	1081	GACATCCGCGACGCAATTTGCCATCGACTACGACCGCTTAGAGGGCTATGTCTACTGGAC	1140
QY	1222	GATGACGAGGTGCGGGCCATCCGCGAGGCGTACCTTGGACGCGTCTGGGCGCGAGACGCTG	1281
DB	1141	GATGACGAGGTGCGGGCCATCCGCGAGGCGTACCTTGGACGCGTCTGGGCGCGAGACGCTG	1200
QY	1282	GTCAACACCGAGATCAACGACCCCGATGGCATTCGCGGTGCACTGGGTGGCCCGAAACCTC	1341
DB	1201	GTCAACACCGAGATCAACGACCCCGATGGCATTCGCGGTGCACTGGGTGGCCCGAAACCTC	1260
QY	1342	TACTGGACCGACACGCGGCGACCGCATCGAGGTGACGCGCTCAAGGCGACCTCCCGC	1401
DB	1261	TACTGGACCGACACGCGGCGACCGCATCGAGGTGACGCGCTCAAGGCGACCTCCCGC	1320
QY	1402	AAGATCTGTGTTCGAGGACCTGGACGAGCCCGAGCGCATTCGCACTGACACCCGATGATG	1461
DB	1321	AAGATCTGTGTTCGAGGACCTGGACGAGCCCGAGCGCATTCGCACTGACACCCGATGATG	1380
QY	1462	GGCTCATGTACTGGACAGACTGGGAGAGAACCTTAATTCGAGTGTGCGCACTTGGAT	1521
DB	1381	GGCTCATGTACTGGACAGACTGGGAGAGAACCTTAATTCGAGTGTGCGCACTTGGAT	1440
QY	1522	GGGACGAGCGCGGTGTGCTGCTCAATGCTCTCCCTCGGGTGGCCCAAGCGGCTGCGCCCTG	1581
DB	1441	GGGACGAGCGCGGTGTGCTGCTCAATGCTCTCCCTCGGGTGGCCCAAGCGGCTGCGCCCTG	1500
QY	1582	GACCTGAGGAGGGAAGCTTACTGGGAGACGCGCAAGACAGACAAGATCAGAGTATC	1641
DB	1501	GACCTGAGGAGGGAAGCTTACTGGGAGACGCGCAAGACAGACAAGATCAGAGTATC	1560
QY	1642	AATGTTGATGGGAGGAGCGGACCTCTCTGGAGGACAGCTCCCGCAATTTTGGG	1701
DB	1561	AATGTTGATGGGAGGAGCGGACCTCTCTGGAGGACAGCTCCCGCAATTTTGGG	1620
QY	1702	TTACGCTGTCTGGGGGACTTCTACTTGGACTGCTGGAGCGCGCCAGCATCGAGCGG	1761
DB	1621	TTACGCTGTCTGGGGGACTTCTACTTGGACTGCTGGAGCGCGCCAGCATCGAGCGG	1680
QY	1762	GTGCAAGGTCAGAGCGCGGAGCGTCAATTCATTCAGACGCTGCGCGAGCTGATGGG	1821
DB	1681	GTGCAAGGTCAGAGCGCGGAGCGTCAATTCATTCAGACGCTGCGCGAGCTGATGGG	1740

QY 322 TACCTGAACACAGACGGGGCGCCGCTGACAGACGTGGTCACTCTCGGCTGGTCTCTCCC 381
DB |||||
QY 241 TACCTGAACACAGACGGGGCGCCGCTGACAGACGTGGTCACTCTCGGCTGGTCTCTCCC 300
DB |||||
QY 382 GACGGCTTCGCTGCGACTGGGTGGGCAAGAGCTGTACTGGAACGATCTGAGACCAAC 441
DB |||||
QY 301 GACGGCTTCGCTGCGACTGGGTGGGCAAGAGCTGTACTGGAACGATCTGAGACCAAC 360
DB |||||
QY 442 CGCATCGAGGTGGCCAACTCAATGGCACATCCCGAAGGTGCTCTTCTGCGACGACCTT 501
DB |||||
QY 361 CGCATCGAGGTGGCCAACTCAATGGCACATCCCGAAGGTGCTCTTCTGCGACGACCTT 420
DB |||||
QY 502 GACCAAGCTTAGGGCCCATCGCTTGGACCCCGCTCACGGGTACATGTACTGGAACGACTGG 561
DB |||||
QY 421 GACCAAGCTTAGGGCCCATCGCTTGGACCCCGCTCACGGGTACATGTACTGGAACGACTGG 480
DB |||||
QY 562 GGTGAGACGCCCGGATTTGAGCGGCGAGGGATGGATGGCAGACCCCGAAGATCATTTGTG 621
DB |||||
QY 481 GGTGAGACGCCCGGATTTGAGCGGCGAGGGATGGATGGCAGACCCCGAAGATCATTTGTG 540
DB |||||
QY 622 GACTCGGACATTTACTGCGCCCAATGGACTGACCATCGACCTGGAAGGACGAGACTCTAC 681
DB |||||
QY 541 GACTCGGACATTTACTGCGCCCAATGGACTGACCATCGACCTGGAAGGACGAGACTCTAC 600
DB |||||
QY 682 TGGGCTGACGGCAAGCTCAGCTTTCATCCACCGTGGCAACCTTGGACGGCTCGTTCCGGCAG 741
DB |||||
QY 601 TGGGCTGACGGCAAGCTCAGCTTTCATCCACCGTGGCAACCTTGGACGGCTCGTTCCGGCAG 660
DB |||||
QY 742 AAGGTGTGGAGGGGAGGCTGACGCAACCCCTTGGCCCTTGAAGCTCTCCGGGGACACTCTG 801
DB |||||
QY 661 AAGGTGTGGAGGGGAGGCTGACGCAACCCCTTGGCCCTTGAAGCTCTCCGGGGACACTCTG 720
DB |||||
QY 802 TACTGGAACAGACTGACAGACCGCTTCCATCCATGCACTGCAACAGCGCACTGGGGGAG 861
DB |||||
QY 721 TACTGGAACAGACTGACAGACCGCTTCCATCCATGCACTGCAACAGCGCACTGGGGGAG 780
DB |||||
QY 862 AGGAAGAGAGACTCTGAGTGGCCCTTACTCACCCATGGACATCCAGGTGCTGAGCCAGGAG 921
DB |||||
QY 781 AGGAAGAGAGACTCTGAGTGGCCCTTACTCACCCATGGACATCCAGGTGCTGAGCCAGGAG 840
DB |||||
QY 922 CGGACAGCTTTCTTCCACACTCGCTGTGAGGAGCAATGGCGGTGTCTCCACACTGTGC 981
DB |||||
QY 841 CGGACAGCTTTCTTCCACACTCGCTGTGAGGAGCAATGGCGGTGTCTCCACACTGTGC 900
DB |||||
QY 982 CTGCTGTCCCAAGCGAGCTTTCTACACATGCGCTGCCCAAGGGGTGTCAGCTGCAG 1041
DB |||||
QY 901 CTGCTGTCCCAAGCGAGCTTTCTACACATGCGCTGCCCAAGGGGTGTCAGCTGCAG 960
DB |||||
QY 1042 GACAAACGACGAGAGCTGTAAGGACAGGAGCGGAGGAGTGTGCTGTGCGCCCGCGGACG 1101
DB |||||
QY 961 GACAAACGACGAGAGCTGTAAGGACAGGAGCGGAGGAGTGTGCTGTGCGCCCGCGGAGC 1020
DB |||||
QY 1102 GACCTACGAGAGATCTCGCTGGAACACGCGGAGCTTACCGACATCGTGTGAGAGTGAGC 1161
DB |||||
QY 1021 GACCTACGAGAGATCTCGCTGGAACACGCGGAGCTTACCGACATCGTGTGAGAGTGAGC 1080
DB |||||
QY 1162 GACATCGGACGAGCTTGGCCATCGACTACGACCGCTAGAGGGGTATGTCTACTGGACA 1221
DB |||||
QY 1081 GACATCGGACGAGCTTGGCCATCGACTACGACCGCTAGAGGGGTATGTCTACTGGACA 1140
DB |||||
QY 1222 GATGACGAGGTGCGGGCCATCCGACGAGGCTTACCTGACCGGTCTGGGGCGCAGACGCTG 1281
DB |||||
QY 1141 GATGACGAGGTGCGGGCCATCCGACGAGGCTTACCTGACCGGTCTGGGGCGCAGACGCTG 1200
DB |||||
QY 1282 GTCAAACCCGAGATCAACGACCCCGATGCGGTGCTGACTGGGTGGCCCGAAACCTC 1341
DB |||||
QY 1201 GTCAAACCCGAGATCAACGACCCCGATGCGGTGCTGACTGGGTGGCCCGAAACCTC 1260
DB |||||
QY 1342 TACTGGAACGACGAGGACGAGCGGATCGAGGTGAGCGGCTTCAAGCGACCTTCCCGC 1401
DB |||||
QY 1261 TACTGGAACGACGAGGACGAGCGGATCGAGGTGAGCGGCTTCAAGCGACCTTCCCGC 1320
DB |||||

QY 1402 AAGATCCTGTGTTCGAGGACCTTGGACGAGCCCGGACCATCGCACTGCACCCCGTGTATG 1461
DB |||||
QY 1321 AAGATCCTGTGTTCGAGGACCTTGGACGAGCCCGGACCATCGCACTGCACCCCGTGTATG 1380
DB |||||
QY 1462 GGCCTCATGTACTGGAACGACTGGGGAGAGAACCTTAAATCGAGTGTGCCAACTTGGAT 1521
DB |||||
QY 1381 GGCCTCATGTACTGGAACGACTGGGGAGAGAACCTTAAATCGAGTGTGCCAACTTGGAT 1440
DB |||||
QY 1522 GGGCAGAGAGCGGCTGTGTCTCAATGCTCTCCCTCGGGTGGCCCAAGCGGCTGGCCCTG 1581
DB |||||
QY 1441 GGGCAGAGAGCGGCTGTGTCTCAATGCTCTCCCTCGGGTGGCCCAAGCGGCTGGCCCTG 1500
DB |||||
QY 1582 GACCTGACGAGGGGAAAGCTTACTTGGGGAGACGCAAGACAGACAGATCGAGGTGATC 1641
DB |||||
QY 1501 GACCTGACGAGGGGAAAGCTTACTTGGGGAGACGCAAGACAGACAGATCGAGGTGATC 1560
DB |||||
QY 1642 AATGTTGATGGGACGAGAGGCGGACCTCTCTGAGGGAAGAGCTCCCGACATTTTCGG 1701
DB |||||
QY 1561 AATGTTGATGGGACGAGAGGCGGACCTCTCTGAGGGAAGAGCTCCCGACATTTTCGG 1620
DB |||||
QY 1702 TTCAAGCTGTGTGGGGGACTTCTACTTGGGACTGACTGGCAGCGCGCAGCATCGAGCGG 1761
DB |||||
QY 1621 TTCAAGCTGTGTGGGGGACTTCTACTTGGGACTGACTGGCAGCGCGCAGCATCGAGCGG 1680
DB |||||
QY 1762 GTGCACAAGGTCAAGGGCAGCGGGGACGTCAATTCATTCACCAAGCTGCCGACCTGATGGG 1821
DB |||||
QY 1681 GTGCACAAGGTCAAGGGCAGCGGGGACGTCAATTCATTCACCAAGCTGCCGACCTGATGGG 1740
DB |||||
QY 1822 CTCAAAGCTGTGAATGTGGCCAGAGTGTTCGGAACCAACCCCGTGTGCGGACAGGAACCGG 1881
DB |||||
QY 1741 CTCAAAGCTGTGAATGTGGCCAGAGTGTTCGGAACCAACCCCGTGTGCGGACAGGAACCGG 1800
DB |||||
QY 1882 GGGTGCAGCCACCTGTGCTTCTTCCACCCCAAGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1941
DB |||||
QY 1801 GGGTGCAGCCACCTGTGCTTCTTCCACCCCAAGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1860
DB |||||
QY 1942 CTGAGAGCTGTGAGTGAACATGAAGACCTGCACTGCTGCTGAGGCTTCTTGTGTCTTACC 2001
DB |||||
QY 1861 CTGAGAGCTGTGAGTGAACATGAAGACCTGCACTGCTGCTGAGGCTTCTTGTGTCTTACC 1920
DB |||||
QY 2002 AGCAGAGCGGCATCCACAGGATCTCCCTCGAGACCAATTAACAGAGCTGGCCCATCCG 2061
DB |||||
QY 1921 AGCAGAGCGGCATCCACAGGATCTCCCTCGAGACCAATTAACAGAGCTGGCCCATCCG 1980
DB |||||
QY 2062 CTCACGGCGCTCAAGAGGCTCAGCCCTGAGCTTGTGATGTGCCAACCAACACATCTAC 2121
DB |||||
QY 1981 CTCACGGCGCTCAAGAGGCTCAGCCCTGAGCTTTGATGTGCCAACCAACACATCTAC 2040
DB |||||
QY 2122 TGGACAGAGCTCAGCTCAGGACCATCAGCCGCGCTTTCATGAACGGGAGCTCGGTGGAG 2181
DB |||||
QY 2041 TGGACAGAGCTCAGCTCAGGACCATCAGCCGCGCTTTCATGAACGGGAGCTCGGTGGAG 2100
DB |||||
QY 2182 CACGTGTGAGTGTGGCTTGAATACCCGAGGCGATGGCGGTGATGTGATGGGCAAG 2241
DB |||||
QY 2101 CACGTGTGAGTGTGGCTTGAATACCCGAGGCGATGGCGGTGATGTGATGGGCAAG 2160
DB |||||
QY 2242 AACCTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTGGCGGCTCGACGGGAG 2301
DB |||||
QY 2161 AACCTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTGGCGGCTCGACGGGAG 2220
DB |||||
QY 2302 TTCCGGCAAGTCTCTGTGTGGAGGAGCTTGGACCAACCCGAGGTGCTGGCCCTGGATCCC 2361
DB |||||
QY 2221 TTCCGGCAAGTCTCTGTGTGGAGGAGCTTGGACCAACCCGAGGTGCTGGCCCTGGATCCC 2280
DB |||||
QY 2362 ACAGAGGCTTACATCTTACTGGAACGAGTGGGGCGGCAAGCCGAGGATCGTGGGCGCTTC 2421
DB |||||
QY 2281 ACAGAGGCTTACATCTTACTGGAACGAGTGGGGCGGCAAGCCGAGGATCGTGGGCGCTTC 2340
DB |||||
QY 2422 ATGGAACGGACCAACTGCTGAGCTGTGGACAAAGTGGGCGGCGCAAGCTCACC 2481
DB |||||
QY 2341 ATGGAACGGACCAACTGCTGAGCTGTGGACAAAGTGGGCGGCGCAAGCTCACC 2400
DB |||||
QY 2482 ATTGACTACGCTGACGCGCTCTTACTGGACCGACCTTGACACCAACATGATCGAGTGC 2541
DB |||||

|||||
2401 ATTGACTAGCTGACCGAGCCTCTACTGGACCGACTGGACCAACATGATCAGATCG 2460
QY
2542 TCACACATGCTGGGTGAGAGCGGTGCTGATGTCGCGAGTCTCCGACCCGTTCCGT 2601
Db
2461 TCACACATGCTGGGTGAGAGCGGTGCTGATGTCGCGAGTCTCCGACCCGTTCCGT 2520
QY
2602 CTGACGAGTACAGGATATATCTACTGGACAGACTGGAATCTGCACAGCATTTAGCGG 2661
Db
2521 CTGACGAGTACAGGATATATCTACTGGACAGACTGGAATCTGCACAGCATTTAGCGG 2580
QY
2662 GCCGACAAAGACTAGCGGCGGAAACCGCACCTCTCAGGGCCACTGACTTTCGTGATG 2721
Db
2581 GCCGACAAAGACTAGCGGCGGAAACCGCACCTCTCAGGGCCACTGACTTTCGTGATG 2640
QY
2722 GACATCTCTGGTGTTCATCTCTCCGCGCAGGATGCCCTCAATGACTGTATGCAACAAC 2781
Db
2641 GACATCTCTGGTGTTCATCTCTCCGCGCAGGATGCCCTCAATGACTGTATGCAACAAC 2700
QY
2782 GGGCAGTGTGGCAGCTGTGCTTGGCATCCCGGCGGCACCGCTGGCGCTGCGCCTCA 2841
Db
2701 GGGCAGTGTGGCAGCTGTGCTTGGCATCCCGGCGGCACCGCTGGCGCTGCGCCTCA 2760
QY
2842 CACTACACCTTGGACCCGACAGCGGCAACTGCAGCCGCGCCACCACTTCTTGTGTTTC 2901
Db
2761 CACTACACCTTGGACCCGACAGCGGCAACTGCAGCCGCGCCACCACTTCTTGTGTTTC 2820
QY
2902 AGCCAGAAATCTGCCATCAGTCTCGATGATTCGCGACGACAGCAGCCCGGATCTCATC 2961
Db
2821 AGCCAGAAATCTGCCATCAGTCTCGATGATTCGCGACGACAGCAGCCCGGATCTCATC 2880
QY
2962 CTGCCCTCTGACTGAGTCTGAGAACTGCAAGCACTGCACTGACCTGACCTGCAAGTTTC 3021
Db
2881 CTGCCCTCTGACTGAGTCTGAGAACTGCAAGCACTGCACTGACCTGACCTGCAAGTTTC 2940
QY
3022 ATCTACTGGGTGGATGGCGCCAGAACTCAAGCGAGCCAAAGGACGAGCGGACCCAGCCCC 3081
Db
2941 ATCTACTGGGTGGATGGCGCCAGAACTCAAGCGAGCCAAAGGACGAGCGGACCCAGCCCC 3000
QY
3082 TTTGTTTTGACCTCTCTAGCCAGGCGCAAAACCCAGACAGGAGCCCGACCTCAGC 3141
Db
3001 TTTGTTTTGACCTCTCTAGCCAGGCGCAAAACCCAGACAGGAGCCCGACCTCAGC 3060
QY
3142 ATCGACATCTACGCGGACACTGTTCTGAGAGTGTGCTGAGGAGCCCAATACCATCAAGCTC 3201
Db
3061 ATCGACATCTACGCGGACACTGTTCTGAGAGTGTGCTGAGGAGCCCAATACCATCAAGCTC 3120
QY
3202 CACAGGCTGAGCGGGAAGCCATGGGGTGTGCTGCTGGGAGCCGCAAGCCAGG 3261
Db
3121 CACAGGCTGAGCGGGAAGCCATGGGGTGTGCTGCTGGGAGCCGCAAGCCAGG 3180
QY
3262 GCCATCTGCTCAACCGGAGGAGGTACCTGTACTTCAACCAATGACAGGACCGGGCA 3321
Db
3181 GCCATCTGCTCAACCGGAGGAGGTACCTGTACTTCAACCAATGACAGGACCGGGCA 3240
QY
3322 GCCAAGATCGAAACCGGAGCCCTGAGCGGACCGAGCCGAGGTCTCTTCAACCAAGCCGCG 3381
Db
3241 GCCAAGATCGAAACCGGAGCCCTGAGCGGACCGAGCCGAGGTCTCTTCAACCAAGCCGCG 3300
QY
3382 CTCAATCCGCTGTGGCCCTGTGGTGTGGAACAACAACACTGGGCAAGCTGTTCTGGGTGGAC 3441
Db
3301 CTCAATCCGCTGTGGCCCTGTGGTGTGGAACAACAACAACACTGGGCAAGCTGTTCTGGGTGGAC 3360
QY
3442 GCGGACCTGAGCCGATGAGAGTGTGACTGTCTGAGGGGCAACCGCTGACCTGAG 3501
Db
3361 GCGGACCTGAGCCGATGAGAGTGTGACTGTCTGAGGGGCAACCGCTGACCTGAG 3420
QY
3502 GACGCCAATCTGTGAGCTCTGGGCTGACCACTCTTGGCAAGCATCTCTACTGGATC 3561
Db
3421 GACGCCAATCTGTGAGCTCTGGGCTGACCACTCTTGGCAAGCATCTCTACTGGATC 3480
QY
3562 GACGCCAAGCAGATGATCGAGCTGTGAGAGAACCAACCGGGCAAGCGGACTGCG 3621
|||||

Db
3481 GACCGCCAGCAGCAGATGATCGAGGCTGTGAGAGAACCAACCGGGGCAAGCGGACTCGC 3540
QY
3622 ATCCAGGCGCTGTGTGCCCACTCTCATCTGGCATCCATGAGTGGAGAAAGTCAAGCTCGAG 3681
Db
3541 ATCCAGGCGCTGTGTGCCCACTCTCATCTGGCATCCATGAGTGGAGAAAGTCAAGCTCGAG 3600
QY
3682 GAGTTCTCAGCGCCACCCATGTCGCGTGACAAATGGTGGCTGCTCCACATCTGTATTGCC 3741
Db
3601 GAGTTCTCAGCGCCACCCATGTCGCGTGACAAATGGTGGCTGCTCCACATCTGTATTGCC 3660
QY
3742 AAGGTGATGAGACACCAACGCTGTCTATGCCAGTCCACCTCTGCTCTCTGAGAACCTG 3801
Db
3661 AAGGTGATGAGACACCAACGCTGTCTATGCCAGTCCACCTCTGCTCTCTGAGAACCTG 3720
QY
3802 CTGACCTGTGTGAGACCGCCCACTGTCTCCCGGACCAAGTTTGCATGTGSCACAGGGGAG 3861
Db
3721 CTGACCTGTGTGAGACCGCCCACTGTCTCCCGGACCAAGTTTGCATGTGSCACAGGGGAG 3780
QY
3862 ATCGACTGTATCCCGGGGCTGTGGCTGTGACCGCTTCCGAGTGGATGACAGAGC 3921
Db
3781 ATCGACTGTATCCCGGGGCTGTGGCTGTGACCGCTTCCGAGTGGATGACAGAGC 3840
QY
3922 GACGAGAGGCTGTCCCGCTGTCTCGCGCCAGTTCCTCTGCGCGGGCTCAGTGT 3981
Db
3841 GACGAGAGGCTGTCCCGCTGTCTCGCGCCAGTTCCTCTGCGCGGGCTCAGTGT 3900
QY
3982 GTGACCTGTGCGCTGTGCGACGCGGAGGACAGACTGTTCAGGACCGCTCAGAGAGGCG 4041
Db
3901 GTGACCTGTGCGCTGTGCGACGCGGAGGAGAGACTGTTCAGGACCGCTCAGAGAGGCG 3960
QY
4042 GACTGTGAGCGCATCTGCTGCCCAACCAAGTTCGGTGTGCGGCGGCGCAGTGTCTCCTC 4101
Db
3961 GACTGTGAGCGCATCTGCTGCCCAACCAAGTTCGGTGTGCGGCGGCGCAGTGTCTCCTC 4020
QY
4102 ATCAACAGCAGTGTGCTCTTCCCGACTGTATCGACGGCTCCGACGAGCTCATGTGT 4161
Db
4021 ATCAACAGCAGTGTGCTCTTCCCGACTGTATCGACGGCTCCGACGAGCTCATGTGT 4080
QY
4162 GAAATCAGCAGCGCTCTGAGCAGACGCGCGGCGCACAGAGTGCATCGCGGCCGCTC 4221
Db
4081 GAAATCAGCAGCGCTCTGAGCAGACGCGCGGCGCACAGAGTGCATCGCGGCCGCTC 4140
QY
4222 ATTGCGCATCATCTCTCTCTCTGCTATGAGTGTGCTATTTTGTGTCAGCGCGTG 4281
Db
4141 ATTGCGCATCATCTCTCTCTCTCTGCTATGAGTGTGCTATTTTGTGTCAGCGCGTG 4200
QY
4282 GTGTGCGCAGCGCTATGCGGGGCGCAACCGGGCTCTTCCGCAACGAGTATGTTCAGCGGAGC 4341
Db
4201 GTGTGCGCAGCGCTATGCGGGGCGCAACCGGGCTCTTCCGCAACGAGTATGTTCAGCGGAGC 4260
QY
4342 CCGCAGCTGCGCTCAATTTTCAATAGCCCGGGCGGTTCCAGCATGGCCCTTCACAGGC 4401
Db
4261 CCGCAGCTGCGCTCAATTTTCAATAGCCCGGGCGGTTCCAGCATGGCCCTTCACAGGC 4320
QY
4402 ATCCATGTCGGAAGTCCATGATGAGTCCGTGAGCTGTATGGGGGCGCGGGCGGGTG 4461
Db
4321 ATCCATGTCGGAAGTCCATGATGAGTCCGTGAGCTGTATGGGGGCGCGGGCGGGTG 4380
QY
4462 CCCCTGTACGACCGGAACCACTGTCAGAGGGCTCTGTCAGAGAGCTCTGTCAGACGAG 4521
Db
4381 CCCCTGTACGACCGGAACCACTGTCAGAGGGCTCTGTCAGAGAGCTCTGTCAGACGAG 4440
QY
4522 GCGACGCTGTACCGCGCATCTTGAAACCGCGCGCTCTCCCGGCAAGGACCTCTCCTG 4581
Db
4441 GCGACGCTGTACCGCGCATCTTGAAACCGCGCGCTCTCCCGGCAAGGACCTCTCCTG 4500
QY
4582 TACAAATGAGCATGTTCTACTCTTCAAAATTCGCGCCACTGCGAGACCGTACAGGGCC 4641
Db
4501 TACAAATGAGCATGTTCTACTCTTCAAAATTCGCGCCACTGCGAGACCGTACAGGGCC 4560
QY
4642 TACATCATTCGAGGAATGGGCGCCCGGACGAGCGCTCTGAGACCGAGCTGTGTGACAGC 4701
Db
4561 TACATCATTCGAGGAATGGGCGCCCGGACGAGCGCTCTGAGACCGAGCTGTGTGACAGC 4620

QY 4702 GACTACAGCGCAGCGCTGGAAGGCCAGCAAGTACTACCTGGATTGAACTCGACTCA 4761
D 4703 GACTACAGCGCAGCGCTGGAAGGCCAGCAAGTACTACCTGGATTGAACTCGACTCA 4762
QY 4762 GACCCCTATCCACCCCAACCCAGCCGCCACAGCCAGTACCTGTGCGCGGAGAGACAGCTGC 4821
D 4681 GACCCCTATCCACCCCAACCCAGCCGCCACAGCCAGTACCTGTGCGCGGAGAGACAGCTGC 4740
QY 4822 CGCGCTCGCGCGCCAGCGAGAGAGTACTTCCATCTTCCCGCGCCCTCCGTCGCC 4881
D 4741 CGCGCTCGCGCGCCAGCGAGAGAGTACTTCCATCTTCCCGCGCCCTCCGTCGCC 4800
QY 4882 TGCACGGACTCATCTGACCTCGCGCGGCCACCTCTGCTGCTCTCTGCTGCTCTGTAATA 4941
D 4801 TGCACGGACTCATCTGACCTCGCGCGGCCACCTCTGCTGCTCTCTGCTGCTCTGTAATA 4860
QY 4942 GTTTTAAATATGAACAAAGAAAAATATATTTTATGATTTTAAAAATAAATAAATTTGG 5001
D 4861 GTTTTAAATATGAACAAAGAAAAATATATTTTATGATTTTAAAAATAAATAAATTTGG 4920
QY 5002 GATTTTAAAAACATGAGAAATGTGAACCTGTGAGTGGGTGGCAGCGCTGGGAGAACTTTG 5061
D 4921 GATTTTAAAAACATGAGAAATGTGAACCTGTGAGTGGGTGGCAGCGCTGGGAGAACTTTG 4980
QY 5062 TA 5063
D 4981 TA 4982

RESULT 10

US-09-060-299-31
Sequence 31, Application US/09060299
Patent No. 6545137
GENERAL INFORMATION:
APPLICANT: Todd, John A
APPLICANT: Hess, John W
APPLICANT: Caskey, Charles T
APPLICANT: Cox, Roger D
APPLICANT: Gerhold, David
APPLICANT: Hammond, Holly
APPLICANT: Hey, Patricia
APPLICANT: Kawaguchi, Yoshihiko
APPLICANT: Merriman, Tony R
APPLICANT: Metzker, Michael L
TITLE OF INVENTION: No. 6545137el Receptor
NUMBER OF SEQUENCES: 455
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon and Vanderhye
STREET: 1100 No. 6545137th Glebe Road, Eighth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: US
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/060,299
FILING DATE: 15-APR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/043,553
FILING DATE: 15-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/048,740
FILING DATE: 05-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: B.J.Sadoff
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 620-35
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)816-4091
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 5263 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-060-299-31

Query Match 96.7%; Score 4894.4; DB 4; Length 5263;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 4920; Conservative 0; Mismatches 6; Indels 2; Gaps 2;
QY 136 GCCCGCGGCTCGCGCTCTCTGCTATTTTGCCTAAACCGCGGGACGTACGGCTGTGTGACGCC 195
D 298 GCCCGAGGCTCGCGCTCTCTGCTATTTTGCCTAAACCGCGGGACGTACGGCTGTGTGACGCC 357
QY 196 GCGGAGGTCAAGCTGGAGTCCACCATCGTGGTTCAGCGGCTTGGAGGATGCGGCGCAGTG 255
D 358 GCGGAGGTCAAGCTGGAGTCCACCATCGTGGTTCAGCGGCTTGGAGGATGCGGCGCAGTG 417
QY 256 GACTTCAGTTTCCAAAGGAGCGGTGTACTGGACAGACGTGAGGAGAGGCCATCAAG 315
D 418 GACTTCAGTTTCCAAAGGAGCGGTGTACTGGACAGACGTGAGGAGAGGCCATCAAG 477
QY 316 CAGACCTACCTGAACACAGAGCGGGGCGCGTGCAGAACGTGTCTATCTCCGGCTGTGTC 375
D 478 CAGACCTACCTGAACACAGAGCGGGGCGCGTGCAGAACGTGTGTCTCTCGGCTGTGTC 537
QY 376 TCTCCCGACGGCTCGCTCGACTGGGTGGGCAAGAGCTGTACTTGGACGGACTCAGAG 435
D 538 TCTCCCGACGGCTCGCTCGACTGGGTGGGCAAGAGCTGTACTTGGACGGACTCAGAG 597
QY 436 ACCAAACCGCATCGAGGTGGCCAACTCAATGGCAATCCCGGAAGGTGCTCTTCTGGGAG 495
D 598 ACCAAACCGCATCGAGGTGGCCAACTCAATGGCAATCCCGGAAGGTGCTCTTCTGGGAG 657
QY 496 GACCTTGACACCGCTAGCGGCGCATCGCTTGGACCCCGCTACAGGCTACATGTCTGGACA 555
D 658 GACCTTGACACCGCGAGGGCCATCGCTTGGACCCCGCTACAGGCTACATGTCTGGACA 717
QY 556 GACTGGGGTGAAGACGCCCGGATTTAGCGGGCAGGGATGGATGGCAGCACCCCGGAAGATC 615
D 718 GACTGGGGTGAAGACGCCCGGATTTAGCGGGCAGGGATGGATGGCAGCACCCCGGAAGATC 777
QY 616 ATTGTGACTCGGACATTTACTGGCCCAATGGATGACCATCGAGCTGGAGAGAGCAGAAG 675
D 778 ATTGTGACTCGGACATTTACTGGCCCAATGGATGACCATCGAGCTGGAGAGAGCAGAAG 837
QY 676 CTCTACTGGGCTGACGCCCAAGCTCAGCTTTCATCCACCGTGCACACCTGCGGCTCGTTC 735
D 838 CTCTACTGGGCTGACGCCCAAGCTCAGCTTTCATCCACCGTGCACACCTGCGGCTCGTTC 897
QY 736 CGGCAAGAGTGGTGGAGGGCAGCTGACGACACCCCTTTCGCCCTTTCGCCGGGAC 795
D 898 CGGCAAGAGTGGTGGAGGGCAGCTGACGACACCCCTTTCGCCCTTTCGCCGGGAC 957
QY 796 ACTCTGTACTGGACAGACTGGCAGACCGCTTCATTCATGCTGCAACAGGGCAGCTGGG 855
D 958 ACTCTGTACTGGACAGACTGGCAGACCGCTTCATTCATGCTGCAACAGGGCAGCTGGG 1017
QY 856 GGGAGAGGAGGAGATCTTGTAGTGGCTTCTACTCACCCATGGACATCCAGGTGCTGAGC 915
D 1018 GGGAGAGGAGGAGATCTTGTAGTGGCTTCTACTCACCCATGGACATCCAGGTGCTGAGC 1077
QY 916 CAGGAGCGGACGCTTTCTTCCACACTCGCTGTGAGGAGGACAATGGCGGCTGTCTCCAC 975
D 1078 CAGGAGCGGACGCTTTCTTCCACACTCGCTGTGAGGAGGACAATGGCGGCTGTCTCCAC 1137
QY 976 CTGTGCTGTGCTCCCAAGAGCGCTTTTCTACACATGCGGCTGTGCGGCTGTGCGAG 1035
D 1138 CTGTGCTGTGCTCCCAAGAGCGCTTTTCTACACATGCGGCTGTGCGGCTGTGCGAG 1197

1036 CTGCAGGCAACCGGACAGGACGTGTAAAGCAGAGACCGGAGGAGTCTGCTGCTGCCCGG 1095
1198 CTGCAGGCAACCGGACAGGACGTGTAAAGCAGAGACCGGAGGAGTCTGCTGCTGCCCGG 1257
1096 CGGACGGACCTACGGAGGATCTCGCTGGACACGCGCGGACTTCACCGACATCGTGTGCGAG 1155
1258 CGGACGGACCTACGGAGGATCTCGCTGGACACGCGCGGACTTCACCGACATCGTGTGCGAG 1317
1156 GTGGACGACATCCGGACACGCCATTTGCCATCGACTAGACCCCGCTAGAGGGCTATGCTTAC 1215
1318 GTGGACGACATCCGGACACGCCATTTGCCATCGACTAGACCCCGCTAGAGGGCTATGCTTAC 1377
1216 TGGACAGATGACGAGGTGCGGSCCATCCGCGAGGCGTACTGGACGGGTCTGGGCGCGAG 1275
1378 TGGACAGATGACGAGGTGCGGSCCATCCGCGAGGCGTACTGGACGGGTCTGGGCGCGAG 1437
1276 AGCTGGTCAACACCGAGATCAACACCCCGATCGCGTTCGACTGGGTGCGCCCGA 1335
1438 AGCTGGTCAACACCGAGATCAACACCCCGATCGCGTTCGACTGGGTGCGCCCGA 1497
1336 AACCTTACTGGACCGACACCGGCGACCGGACCGCATCGAGGTGACCGCGCTCAACGGCAC 1395
1498 AACCTTACTGGACCGACACCGGCGACCGGACCGCATCGAGGTGACCGCGCTCAACGGCAC 1557
1396 TCCCGCAAGATCCTGGTGTGCGAGGACCTGGACGAGCCCGGAGCCATCGCACTGCAACCCC 1455
1558 TCCCGCAAGATCCTGGTGTGCGAGGACCTGGACGAGCCCGGAGCCATCGCACTGCAACCCC 1617
1456 GTGATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAAATCGAGTGTGCCAAC 1515
1618 GTGATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAAATCGAGTGTGCCAAC 1677
1516 TTGGATGGGACGAGCGCGTGTGCTGTCAATGCTCTCCCTGGGTGGCCCAACGGCCCTG 1575
1678 TTGGATGGGACGAGCGCGTGTGCTGTCAATGCTCTCCCTGGGTGGCCCAACGGCCCTG 1737
1576 GCCCTGGACCTGACGAGGGAAGCTCTACTGGGAGAGCGCCAGACAGACAGATCGAG 1635
1738 GCCCTGGACCTGACGAGGGAAGCTCTACTGGGAGAGCGCCAGACAGACAGATCGAG 1797
1636 GTGATCAATGTGTGAGGACGAAGAGCGGACCTCTCTGGAGGACAACTCCCGCACATT 1695
1798 GTGATCAATGTGTGAGGACGAAGAGCGGACCTCTCTGGAGGACAACTCCCGCACATT 1857
1696 TTCGGTTTCACTGCTGGGGAATTCTACTTACTGGACTGTGCGACGCGCGCAGCATC 1755
1858 TTCGGTTTCACTGCTGGGGAATTCTACTTACTGGACTGTGCGACGCGCGCAGCATC 1917
1756 GAGCGGGTGCAACAGGTCAAGCGCAGCGGACGTCATCTTGAACAGCTGCGGACCTG 1815
1918 GAGCGGGTGCAACAGGTCAAGCGCAGCGGACGTCATCTTGAACAGCTGCGGACCTG 1977
1816 ATGGGGCTCAAAGCTGTGAATGTGGCCAAAGTTCGTGGAAACCAACCCGTGTGGGACAGG 1875
1978 ATGGGGCTCAAAGCTGTGAATGTGGCCAAAGTTCGTGGAAACCAACCCGTGTGGGACAGG 2037
1876 AACGGGGGTGAGCCACTGTGCTTCTTCAACCCCGACGAAACCCGCTGTGGTGGCCCC 1935
2038 AACGGGGGGTG-AGCCACTGTGCTTCTTCAACCCCGACGAAACCCGCTGTGGTGGCCCC 2096
1936 ATCGGCTTGAAGCTGTGAGTGACATGAAGACCTGTCATCTGCTGAGGCGCTTCTTGGTC 1995
2097 ATCGGCTTGAAGCTGTGAGTGACATGAAGACCTGTCATCTGCTGAGGCGCTTCTTGGTC 2156
1996 TTCACGACGAGCGGCCATCCACAGGATCTCCCTCGAGACCAATTAACACAGCTGGCC 2055
2157 TTCACGACGAGCGGCCATCCACAGGATCTCCCTCGAGACCAATTAACACAGCTGGCC 2216
2056 ATCCGCTCAGGGGCTCAAGGAGGCTCAGCGCTTGGACTTTGATGTGCTCAACACAC 2115
2217 ATCCGCTCAGGGGCTCAAGGAGGCTCAGCGCTTGGACTTTGATGTGCTCAACACAC 2276

2116 ATCTACTGGACAGAGCTCAGCTGAGACCATCAGCGCGCTTTCATGAACGGGAGCTCG 2175
2277 ATCTACTGGACAGAGCTCAGCTGAGACCATCAGCGCGCTTTCATGAACGGGAGCTCG 2336
2176 GTGAGGACAGCTGTGGAGTGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 2235
2337 GTGAGGACAGCTGTGGAGTGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 2396
2236 GGCAGAACTCTTACTGTGGCGGACACTGGGACCAACAGATCGAAGTGGCGGGTGGAC 2295
2397 GGCAGAACTCTTACTGTGGCGGACACTGGGACCAACAGATCGAAGTGGCGGGTGGAC 2456
2296 GGCAGTTCGCGCAAGTCTCGTGTGGAGGACTTGGACAAACCGAGGTGCTGGCCCTG 2355
2457 GGCAGTTCGCGCAAGTCTCGTGTGGAGGACTTGGACAAACCGAGGTGCTGGCCCTG 2516
2356 GATCCCAACAAAGGCTACATCTACTGGACCGAGTGGGCGGCAAGCCGAGGATCGTGGG 2415
2517 GATCCCAACAAAGGCTACATCTACTGGACCGAGTGGGCGGCAAGCCGAGGATCGTGGG 2576
2416 GCCTTCATGGACGGGACCAACTGATGACGCTGGTGGACAAAGTGGGCGGGGCGCAACGAC 2475
2577 GCCTTCATGGACGGGACCAACTGATGACGCTGGTGGACAAAGTGGGCGGGGCGCAACGAC 2636
2476 CTCACCATTTGACTACGCTGACCGCTCTACTGGACCGGACTTGGACCAACATGATC 2535
2637 CTCACCATTTGACTACGCTGACCGCTCTACTGGACCGGACTTGGACCAACATGATC 2696
2536 GAGTCGTCAACATGCTGGTTCAGGAGCGGCTCGTGTGGTGGCGGACTTCCCGCACCGG 2595
2697 GAGTCGTCAACATGCTGGTTCAGGAGCGGCTCGTGTGGTGGCGGACTTCCCGCACCGG 2756
2596 TTCCGCTGACGAGTACAGCGATTTATCTACTGGACAGACTGGAATCTGCAACGAT 2655
2757 TTCCGCTGACGAGTACAGCGATTTATCTACTGGACAGACTGGAATCTGCAACGAT 2816
2656 GAGCGGCGGACAAAGCTAGCGGCGGAAACCGCACCTCTATCCAGGGCACCTGAGCTTC 2715
2817 GAGCGGCGGACAAAGCTAGCGGCGGAAACCGCACCTCTATCCAGGGCACCTGAGCTTC 2876
2716 GTGATGACATCTCTGTTTCCACTCTCCCGCAGGATGGCTCAATGATGTATGAC 2775
2877 GTGATGACATCTCTGTTTCCACTCTCCCGCAGGATGGCTCAATGATGTATGAC 2936
2776 AACAAACGGGACAGTGTGGGACGCTGTCTTCCATCCCGCGGCGGACCGCTGGCGCTG 2835
2937 AACAAACGGGACAGTGTGGGACGCTGTCTTCCATCCCGCGGCGGACCGCTGGCGCTG 2996
2836 GCCTCACACTACACCTTGGACCCCGACGAGCGGCAACTGACAGCCCGCCACCACTTCTTG 2895
2997 GCCTCACACTACACCTTGGACCCCGACGAGCGGCAACTGACAGCCCGCCACCACTTCTTG 3056
2896 CTGTTTACGACAGAAATCTGCCATCAGTGGATGATCCCGGACGACCAAGCAAGCCCGAT 2955
3057 CTGTTTACGACAGAAATCTGCCATCAGTGGATGATCCCGGACGACCAAGCAAGCCCGAT 3116
2956 CTGATCTGCGCTGACGAGTGTGGGACGCTGCTTCCATCCCGCGGCGGACCGCTGGCGCTG 3015
3117 CTGATCTGCGCTGACGAGTGTGGGACGCTGCTTCCATCCCGCGGCGGACCGCTGGCGCTG 3176
3016 AAGTTTCACTACTGGGTGGGCGGCGGCAACATCAAGCGAGGCAAGGACGAGGGGAC 3075
3177 AAGTTTCACTACTGGGTGGGCGGCGGCAACATCAAGCGAGGCAAGGACGAGGGGAC 3236
3076 CAGCCCTTTGTTTGAACCTCTTCCAGCAAGGCGCAAAACCCAGACAGGACCGCCACGAC 3135
3237 CAGCCCTTTGTTTGAACCTCTTCCAGCAAGGCGCAAAACCCAGACAGGACCGCCACGAC 3296
3136 CTCAGCATCGACATCTACGCGGACACTGTTCTGGAGCTGCGAGGCGCCCAATACCAC 3195
3297 CTCAGCATCGACATCTACGCGGACACTGTTCTGGAGCTGCGAGGCGCCCAATACCAC 3356
3196 AACCTCCACAGGCTGAGCGGCGGAAAGCCATGGGGGTGGTGTGCTGGTGGGACCGCGACAAG 3255

ADDRESSEE: Nixon and Vanderhye
STREET: 1100 No. 655654th Glebe Road, Eighth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: US
ZIP: VA 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,923A
FILING DATE: 14-Feb-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB98/01102
FILING DATE: 15-APR-1998
APPLICATION NUMBER: US 60/043,553
FILING DATE: 15-APR-1997
APPLICATION NUMBER: US 60/048,740
FILING DATE: 05-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: B.J.Sadoff
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 620-81
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4091
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 5263 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-402-923A-31

Query Match 96.7%; Score 4894.4; DB 4; Length 5263;

Best Local Similarity 99.8%; Pred. No. 0;
Matches 4920; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

Qy	136	GC	CGCGGCTCGCGCTCTGCTATTGCGCAACCGCGGACGTACGGCTGGTGACGCC	195
Db	298	GC	CGCAGCTCGCGCTCTGCTATTGCGCAACCGCGGACGTACGGCTGGTGACGCC	357
Qy	196	GC	CGGAGTCAAGTGGAGTCCACATCGTGGTACGCGCTGGAGGATGCGCGCAGTG	255
Db	358	GC	CGGAGTCAAGTGGAGTCCACATCGTGGTACGCGCTGGAGGATGCGCGCAGTG	417
Qy	256	GAC	TTCCAGTTCCTCAAGGGAGCGGTGCTGAGCAGACGCTGAGCGAGGCGCCATCAAG	315
Db	418	GAC	TTCCAGTTCCTCAAGGGAGCGGTGCTGAGCAGACGCTGAGCGAGGCGCCATCAAG	477
Qy	316	CAG	ACTACCTGAAACAGACGCGGGCGCGCTGAGCAACGCTGCTCCTCGGCTGGTC	375
Db	478	CAG	ACTACCTGAAACAGACGCGGGCGCGCTGAGCAACGCTGCTCCTCGGCTGGTC	537
Qy	376	TCT	CCGACGCGCTCGCTCGAGTGGGTGGGCAAGAGCTGTACTGAGCGGCTCAGAG	435
Db	538	TCT	CCGACGCGCTCGCTCGAGTGGGTGGGCAAGAGCTGTACTGAGCGGCTCAGAG	597
Qy	436	ACCA	ACCGCATCGAGTGGCGCAACCTCAATGGCAGATCCCGGAAGGTGCTCTCTGGCAG	495
Db	598	ACCA	ACCGCATCGAGTGGCGCAACCTCAATGGCAGATCCCGGAAGGTGCTCTCTGGCAG	657
Qy	496	GAC	CTTGACAGCTAGGGCCATCGCTTGGACCGCGCTCAGCGGTACATGCTACGACA	555
Db	658	GAC	CTTGACAGCTAGGGCCATCGCTTGGACCGCGCTCAGCGGTACATGCTACGACA	717
Qy	556	GAC	TGGGTGAGCGCCCGGATTTGAGCGGACGGGATGGATGGCAGGACCGCGAGATC	615
Db	718	GAC	TGGGTGAGCGCCCGGATTTGAGCGGACGGGATGGATGGCAGGACCGCGAGATC	777

Qy	616	ATT	TGTGAGCTCGGACATTTTACTGGCCCAATGGACTGACCATCGACCTCGAGAGCAGAAG	675
Db	778	ATT	TGTGAGCTCGGACATTTTACTGGCCCAATGGACTGACCATCGACCTCGAGAGCAGAAG	837
Qy	676	CT	CTACTGGGCTGACGCGCAAGCTCAGCTTCACTCACCGTGCACCACTCGAGCGGCTGTC	735
Db	838	CT	CTACTGGGCTGACGCGCAAGCTCAGCTTCACTCACCGTGCACCACTCGAGCGGCTGTC	897
Qy	736	CGG	CAGAAAGTGGTGGAGGGCAGCTGACGACCCCTTCGCCCTGACGCTCTCCGGGAC	795
Db	898	CGG	CAGAAAGTGGTGGAGGGCAGCTGACGACCCCTTCGCCCTGACGCTCTCCGGGAC	957
Qy	796	ACT	CTGTACTGACAGACTGACAGCCCGCTGCATCTCATGCTGCAACAGCGACTGGG	855
Db	958	ACT	CTGTACTGACAGACTGACAGCCCGCTGCATCTCATGCTGCAACAGCGACTGGG	1017
Qy	856	GGG	AGAGAGAGGATCCTGAGTGCCCTTACTACCATGGACATCAGGTGCTGAGC	915
Db	1018	GGG	AGAGAGAGGATCCTGAGTGCCCTTACTACCATGGACATCAGGTGCTGAGC	1077
Qy	916	CAG	AGCGGACGCTTTCTTCCACATGCTGTGAGGAGCAATGGCGGCTGTCTCCAC	975
Db	1078	CAG	AGCGGACGCTTTCTTCCACATGCTGTGAGGAGCAATGGCGGCTGTCTCCAC	1137
Qy	976	CT	GTGCTGTGTCGCCCAAGCGAGCTTTTACACATGCGCTGCCCGACCGGCTGTCAG	1035
Db	1138	CT	GTGCTGTGTCGCCCAAGCGAGCTTTTACACATGCGCTGCCCGACCGGCTGTCAG	1197
Qy	1036	CT	GAGGACAAACGCGAGCGTGAAGGACGAGCGCGGAGGAGTGTCTCTGTCGCCCGG	1095
Db	1198	CT	GAGGACAAACGCGAGCGTGAAGGACGAGCGCGGAGGAGTGTCTCTGTCGCCCGG	1257
Qy	1096	CGA	CGGACCTTACGAGGATCTCGCTGGACACCGCGGACTTACCGACATCGCTGTCAG	1155
Db	1258	CGA	CGGACCTTACGAGGATCTCGCTGGACACCGCGGACTTACCGACATCGCTGTCAG	1317
Qy	1156	GT	GACGACATCCCGGACGCTTGGCCATCGACTACGACCGCTAGAGGCTATGTCTAC	1215
Db	1318	GT	GACGACATCCCGGACGCTTGGCCATCGACTACGACCGCTAGAGGCTATGTCTAC	1377
Qy	1216	TGG	ACAGATGACGAGGTGCGGGCCATCCGAGGCGGTACTCTGAGCGGCTCTGGGCGCAG	1275
Db	1378	TGG	ACAGATGACGAGGTGCGGGCCATCCGAGGCGGTACTCTGAGCGGCTCTGGGCGCAG	1437
Qy	1276	AC	CGTGTGTAACACCGAGATCAACGACCCCGATGCGGTGCGGTGCGGTGCGGCGA	1335
Db	1438	AC	CGTGTGTAACACCGAGATCAACGACCCCGATGCGGTGCGGTGCGGTGCGGCGA	1497
Qy	1336	AAC	CTTACTGGACCGGACGCGGACCGGACCGGATCGAGGTGACGCGCTCAACGGGACC	1395
Db	1498	AAC	CTTACTGGACCGGACGCGGACCGGACCGGATCGAGGTGACGCGCTCAACGGGACC	1557
Qy	1396	TCC	CGCAAGATCTCTGTTGTCGAGGACCTTGAACGAGCCCGGAGCCATCGCATCGACCCC	1455
Db	1558	TCC	CGCAAGATCTCTGTTGTCGAGGACCTTGAACGAGCCCGGAGCCATCGCATCGACCCC	1617
Qy	1456	GT	GATGGGCTCATGTACTGACAGACTGGGAGAGAACCTTAATTCAGTGTGCCAAC	1515
Db	1618	GT	GATGGGCTCATGTACTGACAGACTGGGAGAGAACCTTAATTCAGTGTGCCAAC	1677
Qy	1516	TT	GATGGGACGAGCGGCTGTGCTCAATGCTCTCCCTCGGGTGGCCCAACGGGCTG	1575
Db	1678	TT	GATGGGACGAGCGGCTGTGCTCAATGCTCTCCCTCGGGTGGCCCAACGGGCTG	1737
Qy	1576	GC	CTTGACCTGACAGGAGGAGGAGTCTTATGGGAGACCGCAAGACAGACATCGAG	1635
Db	1738	GC	CTTGACCTGACAGGAGGAGGAGTCTTATGGGAGACCGCAAGACAGACATCGAG	1797
Qy	1636	GT	GATCAATCTGTGAGAGGAGCGGACCTCTCTGAGGAGCAAGCTCCCGCAATT	1695
Db	1798	GT	GATCAATCTGTGAGAGGAGCGGACCTCTCTGAGGAGCAAGCTCCCGCAATT	1857
Qy	1696	TT	CGGTTACGCTGCTGGGCGACTTTCATCTACTGAGCTGACTGGCAGCGCGGACATC	1755

Db	1958		TTCCGGTTCA	CGCTGCTGGGGG	CTTTCATCTACTGGA	CTGACTGTCGACGCCGCGAGCATC	1917
Qy	1756	GAGCGGGTGA	CAAGAGTCA	AGGCCAGCGGGAG	CGTTCATTTGATACCA	CGAGCTGCCGACCTG	1815
Db	1918	GAGCGGGTGA	CAAGAGTCA	AGGCCAGCGGGAG	CGTTCATTTGATACCA	CGAGCTGCCGACCTG	1977
Qy	1816	ATGGGGCTCA	AAAGCTGTGA	ATGTGGCCAA	GGTCTGTCGGAACCA	ACCCGTGTGCGGACAGG	1875
Db	1978	ATGGGGCTCA	AAAGCTGTGA	ATGTGGCCAA	GGTCTGTCGGAACCA	ACCCGTGTGCGGACAGG	2037
Qy	1876	AACGGGGGTG	CAGCCACCTGTG	CTTTCACACCCCA	CGGCAACCCCGGTGTG	CGCTGCGCC	1935
Db	2038	AACGGGGGTG	-AGGCCACCTGTG	CTTTCACACCCCA	CGGCAACCCCGGTGTG	CGCTGCGCC	2096
Qy	1936	ATCGGCTTGA	AGCTGCTGAGTGA	CAATGAAGACCTG	CAATCGTGCCTGAGGCTCT	CTCTGCGTC	1995
Db	2097	ATCGGCTTGA	AGCTGCTGAGTGA	CAATGAAGACCTG	CAATCGTGCCTGAGGCTCT	CTCTGCGTC	2156
Qy	1996	TTTCAACGAC	GAGCGCCATCC	CAGGATCTTCC	TCGAGACCAATAA	CAACGACGTGGCC	2055
Db	2157	TTTCAACGAC	GAGCGCCATCC	CAGGATCTTCC	TCGAGACCAATAA	CAACGACGTGGCC	2216
Qy	2056	ATCCCGCTCA	CGGCGCTCA	AGGAGGCCCTC	AGCCCTTGACCTTTGATGTG	TCACAAACCCAC	2115
Db	2217	ATCCCGCTCA	CGGCGCTCA	AGGAGGCCCTC	AGCCCTTGACCTTTGATGTG	TCACAAACCCAC	2276
Qy	2116	ATCTACTTGA	CAGACGTCAG	CCCTCAAGACCAT	CAGCCGCGCTTTCATGAACGGGAGCTCG	2175	
Db	2277	ATCTACTTGA	CAGACGTCAG	CCCTCAAGACCAT	CAGCCGCGCTTTCATGAACGGGAGCTCG	2336	
Qy	2176	GTGAGACAG	CTGTGTGGAGTTT	TGGCCTTTGACTAC	CCCGAGGGGCATGGCCGCTTGAGCTGGATG	2235	
Db	2337	GTGAGACAG	CTGTGTGGAGTTT	TGGCCTTTGACTAC	CCCGAGGGGCATGGCCGCTTGAGCTGGATG	2396	
Qy	2236	GGCAAGAA	CCCTCTACTG	CGGCGCA	CACCTGGBGACAA	CAGATTCGGAATGTCGAGCTGGGCTGGAC	2295
Db	2397	GGCAAGAA	CCCTCTACTG	CGGCGCA	CACCTGGBGACAA	CAGATTCGGAATTCGAGCTGGGCTGGAC	2456
Qy	2296	GGGCAGTT	TCGGCAAGTCT	CGTGTGTGAGGGACT	TGCAACCCCGAGGTCGCTGGCCCTG	2355	
Db	2457	GGGCAGTT	TCGGCAAGTCT	CGTGTGTGAGGGACT	TGCAACCCCGAGGTCGCTGGCCCTG	2516	
Qy	2356	GATCCCA	CCAAAGGGCTAC	ACTGACCGAGTGGGGCGGCA	AGCCGAGGATCGTGC	2415	
Db	2517	GATCCCA	CCAAAGGGCTAC	ACTGACCGAGTGGGGCGGCA	AGCCGAGGATCGTGC	2576	
Qy	2416	GCCTTCATGA	CGGGA	CAAACTG	CATGACCGTGTGTGA	CAAGGTGGCCCGGGCCAAACGAC	2475
Db	2577	GCCTTCATGA	CGGGA	CAAACTG	CATGACCGTGTGTGA	CAAGGTGGCCCGGGCCAAACGAC	2636
Qy	2476	CTCACCA	TTGACTAC	CGCTGAC	CGCCCTCTACTG	GACCGACCTTGGACACCAACATGATC	2535
Db	2637	CTCACCA	TTGACTAC	CGCTGAC	CGCCCTCTACTG	GACCGACCTTGGACACCAACATGATC	2696
Qy	2536	GAGTCGT	CTCAA	CATGCTGGGT	CAGGAGCGGTCTGTGATTTGCGGACGATCTCCCGCACCCG	2595	
Db	2697	GAGTCGT	CTCAA	CATGCTGGGT	CAGGAGCGGTCTGTGATTTGCGGACGATCTCCCGCACCCG	2756	
Qy	2596	TTGCGTCTGA	CGCAGTAC	AGCGATTATATCT	ACTTGA	CAGACTGGAAATTCGCAACGATTT	2655
Db	2757	TTGCGTCTGA	CGCAGTAC	AGCGATTATATCT	ACTTGA	CAGACTGGAAATTCGCAACGATTT	2816
Qy	2656	GAGGGGCGCA	CAGACTAG	GGGCGGAA	CGGCACCTCTATCC	AGGGCACCTTGGACTTC	2715
Db	2817	GAGGGGCGCA	CAGACTAG	GGGCGGAA	CGGCACCTCTATCC	AGGGCACCTTGGACTTC	2876
Qy	2716	GTGATGACA	CTCTGGGTG	TTCCACTCT	CTCCGCGCAGGATGGCCTCAATGACTGTATG	TCAC	2775
Db	2877	GTGATGACA	CTCTGGGTG	TTCCACTCT	CTCCGCGCAGGATGGCCTCAATGACTGTATG	TCAC	2936
Qy	2776	AACAA	CGGGCAGTGTGG	GCAGCTGTGCTTTGCCAT	TCCCGCGGCGCACCGCTCGGCTGC	2835	

Db	2937	AAACAACGGGCGAGTGTGGCGCAGCTGTGCTTTGCCATCCCGCGCGCCACACCGCTCGGCGTCG	2999
Qy	2836	GCCTCACACTACACCCCTGGAGCCCGCAGCGCGGCAACTCAGACCCGCGCCACCACTTCTCTTG	2895
Db	2997	GCCTCACACTACACCCCTGGAGCCCGCAGCGCGGCAACTCAGACCCGCGCCACCACTTCTCTTG	3056
Qy	2896	CTGTTTCAGCCAGAAATCTGCCATCAGTTCGGATGATCCCGGACGACCAAGCACGCCCGGAT	2955
Db	3057	CTGTTTCAGCCAGAAATCTGCCATCAGTTCGGATGATCCCGGACGACCAAGCACGCCCGGAT	3116
Qy	2956	CTCATCTGCGCCCTGCATGCAGCTGAGGAAAGTCAAGAGCCATCGACTATGACCCACTTCGGAC	3015
Db	3117	CTCATCTGCGCCCTGCATGCAGCTGAGGAAAGTCAAGAGCCATCGACTATGACCCACTTCGGAC	3176
Qy	3016	AAAGTTCACTACTGGGTGGATGGCGGCCAGAAATCAAGCGAGCCAAAGGACGACGCGGAC	3075
Db	3177	AAAGTTCACTACTGGGTGGATGGCGGCCAGAAATCAAGCGAGCCAAAGGACGACGCGGAC	3236
Qy	3076	CAGCCCTTTGTTTGACCTCTCTGAGCCAAAGGCCAAACCCAGACGAGCCACCAATACCATC	3135
Db	3237	CAGCCCTTTGTTTGACCTCTCTGAGCCAAAGGCCAAACCCAGACGAGCCACCAATACCATC	3296
Qy	3136	CTCAGCATCGACATCTACAGCCGCGACACTGTTCTTGACGCTGCGAGGCGCCACCAATACCATC	3195
Db	3297	CTCAGCATCGACATCTACAGCCGCGACACTGTTCTTGACGCTGCGAGGCGCCACCAATACCATC	3356
Qy	3196	AAAGTCCACAGGCTGAGCGGGGAAGCCATGGGGGTGTGCTGCTGGTGGGACCGCGACAAG	3255
Db	3357	AAAGTCCACAGGCTGAGCGGGGAAGCCATGGGGGTGTGCTGCTGGTGGGACCGCGACAAG	3416
Qy	3256	CCCAGGGCCATCGTGTCTCAACGCGGAGCGAGGTACTGTATTCTCAACCAATGCAGGGAC	3315
Db	3417	CCCAGGGCCATCGTGTCTCAACGCGGAGCGAGGTACTGTATTCTCAACCAATGCAGGGAC	3476
Qy	3316	CGGCGACGCAAGATCGAACAACGCGCAGCCCTGGACGCGACCGAGCGCGAGGTCTCTTTCACC	3375
Db	3477	CGGCGACGCAAGATCGAACAACGCGCAGCCCTGGACGCGACCGAGCGCGAGGTCTCTTTCACC	3536
Qy	3376	ACGGGCTCATCGCCCTGTGGCCCTGTGGTGGACAAACACACTGGGCAAGCTGTTCTGG	3435
Db	3537	ACGGGCTCATCGCCCTGTGGCCCTGTGGTGGTAGACAAACACACTGGGCAAGCTGTTCTGG	3596
Qy	3436	GTGSAACGCGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTTCAGGGGCCAAACCGCTGCACC	3495
Db	3597	GTGSAACGCGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTTCAGGGGCCAAACCGCTGCACC	3656
Qy	3496	CTGAGGAGCGCCAAACATCTGTGACGCTCTGGGCGCTGACCAATCTTGGCAAGCATCTCTAC	3555
Db	3657	CTGAGGAGCGCCAAACATCTGTGACGCTCTGGGCGCTGACCAATCTTGGCAAGCATCTCTAC	3716
Qy	3556	TGGATCGACCGCCAGCAGCAGATGATCGAGCGTGTGAGAAAGACCAACCGGGGACAAGCGG	3615
Db	3717	TGGATCGACCGCCAGCAGCAGATGATCGAGCGTGTGAGAAAGACCAACCGGGGACAAGCGG	3776
Qy	3616	ACTCGCATCCAGGGCCGTGTGCGCCACCTCACTGGCATCCATCAGTGGAGGGAAGTCAGC	3675
Db	3777	ACTCGCATCCAGGGCCGTGTGCGCCACCTCACTGGCATCCATCAGTGGAGGGAAGTCAGC	3836
Qy	3676	CTGAGGAGGTTCTCAGCCCAACCATGTGCCGTGTGACAAATGGTGGCTGTCTCCCAATCTGT	3735
Db	3837	CTGAGGAGGTTCTCAGCCCAACCATGTGCCGTGTGACAAATGGTGGCTGTCTCCCAATCTGT	3896
Qy	3736	ATTGCCAAGGGTGTGAGGACCAACCGGTGCTCATGCCAGTCCACCTCGTGTCTCTGCGAG	3795
Db	3897	ATTGCCAAGGGTGTGAGGACCAACCGGTGCTCATGCCAGTCCACCTCGTGTCTCTGCGAG	3956
Qy	3796	AACCTGCTGACCTGTGAGAGCGCCCGACCTGCTCTCCCGGACCAAGTTTGATGTGCCACA	3855
Db	3957	AACCTGCTGACCTGTGAGAGCGCCCGACCTGCTCTCCCGGACCAAGTTTGATGTGCCACA	4016
Qy	3856	GGGAGATCGATGTATATCCCGGGGCGTGGCGTGTGAACGGCTTTTCCGAGTGCATGAC	3915
Db	4017	GGGAGATCGATGTATATCCCGGGGCGTGGCGTGTGAACGGCTTTTCCGAGTGCATGAC	4076

263 AGTTTTCCAGAGGACCGTGTACTGGACAGACGTGAGCGAGGAGGCCCATCAAGCAGACCT 322
| | | | |
287 AGTTTTCCAGAGGACCGTGTACTGGACAGACGTGAGCGAGGAGGCCCATCAAGCAGACCT 346
| | | | |
323 ACCTGAACAGACGAGGAGCCCGGTGCAAGACGTGGTCAATCTCCGGCCCTGTCTCTCCCG 382
| | | | |
347 ACCTGAACAGACGAGGAGCCCGGTGCAAGACGTGGTCAATCTCCGGCCCTGTCTCTCCCG 406
| | | | |
383 ACAGCCTCGCTGCGACTGGGTGGGCAAGACGTGTACTGACCGGACTCAGAGACCAACC 442
| | | | |
407 ACAGCCTCGCTGCGACTGGGTGGGCAAGACGTGTACTGACCGGACTCAGAGACCAACC 466
| | | | |
443 GCATCGAGGTGGCCAACTCAATGGCACAATCCGGAGGGTCTCTTCTGGCAGACCTTGG 502
| | | | |
467 GCATCGAGGTGGCCAACTCAATGGCACAATCCGGAGGGTCTCTTCTGGCAGACCTTGG 526
| | | | |
503 ACCAGCCTTAGGGCCCATCGCTTGGACCCCGCTCAACGGGTACATGTACTGACAGACTGGG 562
| | | | |
527 ACCAGCAGGGCCCATCGCTTGGACCCCGCTCAACGGGTACATGTACTGACAGACTGGG 586
| | | | |
563 GTGAGACGCCCGGATTTAGCGGGCAGGGATGGATGGCAGCACCGGAGAGATCAATTGTGG 622
| | | | |
587 GTGAGACGCCCGGATTTAGCGGGCAGGGATGGATGGCAGCACCGGAGAGATCAATTGTGG 646
| | | | |
623 ACTCGGACATTTACTTGGCCCAATGGACTGACCATCGACCTGGAGGAGCAGAGACTCTACT 682
| | | | |
647 ACTCGGACATTTACTTGGCCCAATGGACTGACCATCGACCTGGAGGAGCAGAGACTCTACT 706
| | | | |
683 GGGCTGACGCCAAGCTCAGCTTTCATCCACCGTGCACACTTGGACGGGTCTGTTCCGGCAGA 742
| | | | |
707 GGGCTGACGCCAAGCTCAGCTTTCATCCACCGTGCACACTTGGACGGGTCTGTTCCGGCAGA 766
| | | | |
743 AGTGTGTGGAGGACGCTGACGACCGCTTTCGCCCTTGACGCTCTCCGGGAGACATCTGT 802
| | | | |
767 AGTGTGTGGAGGACGCTGACGACCGCTTTCGCCCTTGACGCTCTCCGGGAGACATCTGT 826
| | | | |
803 ACTGACAGACTGTCAGACACCGCTTTCATCCATGCTGCAACAGCGCACTGGGGGGAAGA 862
| | | | |
827 ACTGACAGACTGTCAGACACCGCTTTCATCCATGCTGCAACAGCGCACTGGGGGGAAGA 886
| | | | |
863 GGAAGGAGATCCTGAGTGCCCTTACTCACCCATGGACATCCAGGTGCTGAGCCAGGAGC 922
| | | | |
887 GGAAGGAGATCCTGAGTGCCCTTACTCACCCATGGACATCCAGGTGCTGAGCCAGGAGC 946
| | | | |
923 GGCAGCCTTTCTTCCACACTCGCTGTGAGGAGGACAAATGGCGGCTGCTCCACCTGTGCC 982
| | | | |
947 GGCAGCCTTTCTTCCACACTCGCTGTGAGGAGGACAAATGGCGGCTGCTCCACCTGTGCC 1006
| | | | |
983 TGCTGTCCCAAGCGAGGCTTTCTACACATCGCTGCGCCACCGGTGTGCGAGCTGCAGG 1042
| | | | |
1007 TGCTGTCCCAAGCGAGGCTTTCTACACATCGCTGCGCCACCGGTGTGCGAGCTGCAGG 1066
| | | | |
1043 ACAACGCGAGGACGTGTAAGCGAGGACCGAGGAGGTGCTGCTGCTGGCCCGCGGACGG 1102
| | | | |
1067 ACAACGCGAGGACGTGTAAGCGAGGACCGAGGAGGTGCTGCTGCTGGCCCGCGGACGG 1126
| | | | |
1103 ACCTACGAGGATCTCGTGTGACACGCGGAGCTTCAACGACATCGTGTGCTGAGGTGAGCG 1162
| | | | |
1127 ACCTACGAGGATCTCGTGTGACACGCGGAGCTTTTACCGACATCGTGTGCTGAGGTGAGCG 1186
| | | | |
1163 ACATCCGCGACGCAATTTGCCATTCGACTACGACCCGCTAGAGGGCTATGTCTTACTGGACAG 1222
| | | | |
1187 ACATCCGCGACGCAATTTGCCATTCGACTACGACCCGCTAGAGGGCTATGTCTTACTGGACAG 1246
| | | | |
1223 ATGACGAGGTGCGGGCCATCCGACAGGGCGTACTGGAACGGGTCTGCGGCGCAGACGCTGG 1282
| | | | |
1247 ATGACGAGGTGCGGGCCATCCGACAGGGCGTACTTGGACGGGTCTGCGGCGCAGACGCTGG 1306
| | | | |
1283 TCAACACGAGATCAACGACCCCGATGGCATTCGGGTGCACTGGGTGGCCCGGAAACCTCT 1342
| | | | |
1307 TCAACACGAGATCAACGACCCCGATGGCATTCGGGTGCACTGGGTGGCCCGGAAACCTCT 1366
| | | | |

1343 ACTGACCGACACGCGGACCGGACCGCATTCGAGGTGACGCGCTCAACGCGACCTCCCGCA 1402
| | | | |
1367 ACTGACCGACACGCGGACCGGACCGCATTCGAGGTGACGCGCTCAACGCGACCTCCCGCA 1426
| | | | |
1403 AGATCTCTGGTGTTCGAGGAGCCTTGACGAGCCCGGAGCCATCGCACTGCAACCCCGTGTGG 1462
| | | | |
1427 AGATCTCTGGTGTTCGAGGAGCCTTGACGAGCCCGGAGCCATCGCACTGCAACCCCGTGTGG 1486
| | | | |
1463 GCCTCATGTATTCGACACAGCTGGGGAGAGAACCTTAATCGAGTGTGCCAACTTGGATG 1522
| | | | |
1487 GCCTCATGTATTCGACACAGCTGGGGAGAGAACCTTAATCGAGTGTGCCAACTTGGATG 1546
| | | | |
1523 GGCAGGAGCGCGTGTCTGTCAATGCTCCCTCGGGTGGCCCAACGCGCTGGGCCCTTGG 1582
| | | | |
1547 GGCAGGAGCGCGTGTCTGTCAATGCCCTCCCTCGGGTGGCCCAACGCGCTTGGGCCCTTGG 1606
| | | | |
1583 ACCTGCAAGGAGGGGAAGCTCTTACTGGGGAGACGCAAGACAGACAAGATCGAGGTGTATCA 1642
| | | | |
1607 ACCTGCAAGGAGGGGAAGCTCTTACTGGGGAGACGCAAGACAGACAAGATCGAGGTGTATCA 1666
| | | | |
1643 ATGTTGATGGAGCAAGAGGGGAGACCTCTCTGGAGGACAGCTCCCGCACATTTTCGGGT 1702
| | | | |
1667 ATGTTGATGGAGCAAGAGGGGAGACCTCTCTGGAGGACAGCTCCCGCACATTTTCGGGT 1726
| | | | |
1703 TCACGCTGTCTGGGGGACTTTCATCTACTTGGACTGACTGGCAGCGCGCAGCATCGAGCGGG 1762
| | | | |
1727 TCACGCTGTCTGGGGGACTTTCATCTACTTGGACTGACTGGCAGCGCGCAGCATCGAGCGGG 1786
| | | | |
1763 TGCAACAAGGTCAAGGCGCAGCGGAGCGTTCATCTTACACCAAGCTCCCGACCTGTATGGGGC 1822
| | | | |
1787 TGCAACAAGGTCAAGGCGCAGCGGAGCGTTCATCTTACCAAGCTCCCGACCTGTATGGGGC 1846
| | | | |
1823 TCACAAGTGTGAATGTTCGGCCAAAGTTCGTTCGGAACCAACCCGTGTGCGGACAGAGACGGG 1882
| | | | |
1847 TCAAAAGTGTGAATGTTCGGCCAAAGTTCGTTCGGAACCAACCCGTGTGCGGACAGAGACGGG 1906
| | | | |
1883 GGTGCAGCCACTGTGTCTTTCACACCCACGCAACCCGGTGTGGCTGCGCCCATCGGCC 1942
| | | | |
1907 GGTGCAGCCACTGTGTCTTTCACACCCACGCAACCCGGTGTGGCTGCGCCCATCGGCC 1966
| | | | |
1943 TGGAGCTGTCTGAGTGAACATGAAGACCTGCAATCTGCTGAGGCTTCTTGTGCTTTCACCA 2002
| | | | |
1967 TGGAGCTGTCTGAGTGAACATGAAGACCTGCAATCTGCTGAGGCTTCTTGTGCTTTCACCA 2026
| | | | |
2003 GCAGAGCGGCCATCCACAGGATCTCCCTCGAGACCAATTAACAAGTGTGGCCATCCCGC 2062
| | | | |
2027 GCAGAGCGGCCATCCACAGGAT - TCCCTCGAGACCAATTAACAAGTGTGGCCATCCCGC 2085
| | | | |
2063 TCAGCGCGCTCAAGGAGGCTTCAGCCCTTGGACTTTGATGTGTCCAAACAACACATCTACT 2122
| | | | |
2086 TCAGCGCGCTCAAGGAGGCTTCAGCCCTTGGACTTTGATGTGTCCAAACAACACATCTACT 2145
| | | | |
2123 GGCAGACGCTCAGCCTGAAGACCATCAGCCGCGCTTCATGAACGAGGAGCTCGGTGGAGC 2182
| | | | |
2146 GGCAGACGCTCAGCCTGAAGACCATCAGCCGCGCTTCATGAACGAGGAGCTCGGTGGAGC 2205
| | | | |
2183 ACCTGTTGGAGTTTGGCTTGTACTACCCCGAGGCGATGGCGCTTGAAGTGGGAGC 2242
| | | | |
2206 ACCTGTTGGAGTTTGGCTTGTACTACCCCGAGGCGATGGCGCTTGAAGTGGGAGC 2265
| | | | |
2243 ACCTCTACTTGGGCGGACACTGGGACCAACAGAAATCGAAAGTGGCGGCTGGACGGGACGT 2302
| | | | |
2266 ACCTCTACTTGGGCGGACACTGGGACCAACAGAAATCGAAAGTGGCGGCTGGACGGGACGT 2325
| | | | |
2303 TCCGGCAAGTCTCTGTGTGAGGGAATTGGAACAACCGAGGTCTGTGGCTTGGATCCCA 2362
| | | | |
2326 TCCGGCAAGTCTCTGTGTGAGGGAATTGGAACAACCGAGGTCTGTGGCTTGGATCCCA 2385
| | | | |
2363 CCAGGGCTACATCTACTTGGACCGAGTGGGGCGGCAAGCCGAGGATCGTGGGGCCTTCA 2422
| | | | |
2386 CCAAGGGCTACATCTACTTGGACCGAGTGGGGCGGCAAGCCGAGGATCGTGGGGCCTTCA 2445
| | | | |
2423 TGGACGGGACCAACTGCAATGACGCTGTGGACAAGGTGGGGCGGGCCAAACGACCTCACCA 2482
| | | | |

Db 2446 TGGACGGACCACTGTCATGACGCTGGTGGACAAAGGTGGCCGGGCAACGACCTCAACA 2505
QY 2483 TTGACTACCGCTGACACGACCGCTTACTGACGCGAAGCTGGACACCAACATGATCGAGTCGT 2542
Db 2506 TTGACTACCGCTGACACGACCGCTTACTGACGCGAAGCTGGACACCAACATGATCGAGTCGT 2565
QY 2543 CCAACATGCTGGGTGACGAGCGGGTCTGATTTCCGACGATCTCCGACCGCTTCGGTTC 2602
Db 2566 CCAACATGCTGGGTGACGAGCGGGTCTGATTTCCGACGATCTCCGACCGCTTCGGTTC 2625
QY 2603 TGACGCGATACAGCGATATATATCTACTGACAGACTGGAACTCTGACAGCACTTGAGCGGG 2662
Db 2626 TGACGCGATACAGCGATATATCTACTGACAGACTGGAACTCTGACAGCACTTGAGCGGG 2685
QY 2663 CCGACAAGACTAGCGCGCGGAACCGGACCCCTCATTCAGGGGCCACTGAGTTCGTGATGG 2722
Db 2686 CCGACAAGACTAGCGCGCGGAACCGGACCCCTCATTCAGGGGCCACTGAGTTCGTGATGG 2745
QY 2723 ACATCTGTGTTCACACTCTCCCGCAGGATGGCTCAATGACTGTATGACACAAACG 2782
Db 2746 ACATCTGTGTTCACACTCTCCCGCAGGATGGCTCAATGACTGTATGACACAAACG 2805
QY 2783 GGCAGTGTGGGAGCTGTGCTTGGCATCCCGCGGGCCACCGCTGCGGCTCGCTCAC 2842
Db 2806 GGCAGTGTGGGAGCTGTGCTTGGCATCCCGCGGGCCACCGCTGCGGCTCGCTCAC 2865
QY 2843 ACTACACCTGTGACCCGACGACCGGAACTGACGACCGCCGACCACTTCTGTGTTCA 2902
Db 2866 ACTACACCTGTGACCCGACGACCGGAACTGACGACCGCCGACCACTTCTGTGTTCA 2925
QY 2903 GCCAGAACTGTCCATCAGTCGGATGATCCGGACGACGACGACCGCGGATCTCATCC 2962
Db 2926 GCCAGAACTGTCCATCAGTCGGATGATCCGGACGACGACGACCGCGGATCTCATCC 2985
QY 2963 TGCCCTCTGATGGACTGAGGAACGTCAAAGCCATCGACTATGACCCACTGGACAAAGTTCA 3022
Db 2986 TGCCCTCTGATGGACTGAGGAACGTCAAAGCCATCGACTATGACCCACTGGACAAAGTTCA 3045
QY 3023 TCTACTGGGTGATGGGCGCCAGAACATCAAGCAGGACCAAGGACGAGGACCGCCCT 3082
Db 3046 TCTACTGGGTGATGGGCGCCAGAACATCAAGCAGGACCAAGGACGAGGACCGCCCT 3105
QY 3083 TTGTTTTCACCTCTCTGAGCCAAAGCCAAACCCAGACGAGCGAGCCCAAGACCTCAGCA 3142
Db 3106 TTGTTTTCACCTCTCTGAGCCAAAGCCAAACCCAGACGAGGACCGCCCAAGACCTCAGCA 3165
QY 3143 TCGACATCTACAGCCGACACTGTTCTGAGCGTGGAGGCCAACCAATACCAACGTC 3202
Db 3166 TCGACATCTACAGCCGACACTGTTCTGAGCGTGGAGGCCAACCAATACCAACGTC 3225
QY 3203 ACAGGCTGAGCGGGAAGCCATGGGGTGGTGGTGGGGACCGGACGAGCCAGG 3262
Db 3226 ACAGGCTGAGCGGGAAGCCATGGGGTGGTGGTGGGGACCGGACGAGCCAGG 3285
QY 3263 CCATCGTCTCAACCGCGAGCGAGGTTACCTGTACTTCAACAAATGACGAGCAGGCGAG 3322
Db 3286 CCATCGTCTCAACCGCGAGCGAGGTTACCTGTACTTCAACAAATGACGAGCAGGCGAG 3345
QY 3323 CCAAGATCGAACCGCGACCTTGACGCGACCGAGCGGAGTCTCTTTCACCAACCGGCC 3382
Db 3346 CCAAGATCGAACCGCGACCTTGACGCGACCGAGCGGAGTCTCTTTCACCAACCGGCC 3405
QY 3383 TCATCCGCCCTGTGGCTGTGGTGGACACACACTGGGCAAGCTGTCTGGGTGGACG 3442
Db 3406 TCATCCGCCCTGTGGCTGTGGTGGACACACACTGGGCAAGCTGTCTGGGTGGACG 3465
QY 3443 CGGACCTGAAGCGATTGAGAGCTGTGACTGTGACGGGCGCAACCGGCTGACCCCTGGAGG 3502
Db 3466 CGGACCTGAAGCGATTGAGAGCTGTGACTGTGACGGGCGCAACCGGCTGACCCCTGGAGG 3525
QY 3503 ACGCCAACTGTGAGGCTCTGGGCTGACCATCTTGGCAAGCATCTTACTTGATTCG 3562

Db 3526 AGGCCAACTCGTCGACGCTCTGGGCTTGACCATCTTGGCAAGCATCTTACTCGATCG 3585
QY 3563 ACCGCGACGACGAGATGATCGAGCGTGTGGAGAAAGACACACCGGGGACAAAGCGGACTCGCA 3622
Db 3586 ACCGCGACGACGAGATGATCGAGCGTGTGGAGAAAGACACACCGGGGACAAAGCGGACTCGCA 3645
QY 3623 TCCAGGGCGGTGTGCGCCACCTCTCACTGGCATCCATGACATGAGTGGAGAACTCAGCCTGGAGG 3682
Db 3646 TCCAGGGCGGTGTGCGCCACCTCTCACTGGCATCCATGACATGAGTGGAGAACTCAGCCTGGAGG 3705
QY 3683 AGTTCTCAGCCCAACCATGTGCCGTGACATGTGGTGTCTCCACATCTGTATTGCA 3742
Db 3706 AGTTCTCAGCCCAACCATGTGCCGTGACATGTGGTGTCTCCACATCTGTATTGCA 3765
QY 3743 AGGTTGATGGACACCACTGCTCATGCTCCAGTCCACTCTGCTCTCTGACAACTTCG 3802
Db 3766 AGGTTGATGGACACCACTGCTCATGCTCCAGTCCACTCTGCTCTCTGACAACTTCG 3825
QY 3803 TGACCTGTGGAGAGCGCCCACTGCTCCCGGACCAAGTTTGCATGTGCCACAGGGAGA 3862
Db 3826 TGACCTGTGGAGAGCGCCCACTGCTCCCGGACCAAGTTTGCATGTGCCACAGGGAGA 3885
QY 3863 TCGACTGTATCCCGGGGCTTGGCGCTGTGACGGCTTTCGCGAGTGCATGACACAGAGCG 3922
Db 3886 TCGACTGTATCCCGGGGCTTGGCGCTGTGACGGCTTTCGCGAGTGCATGACACAGAGCG 3945
QY 3923 ACGAGGAGGTGCTCCCGTGTGCTCCGCGGCCAGTTCCCTCTGCGCGGGGTGAGTGTG 3982
Db 3946 ACGAGGAGGTGCTCCCGTGTGCTCCGCGGCCAGTTCCCTCTGCGCGGGGTGAGTGTG 4005
QY 3983 TGGACCTGCGCTGCGGTGCGACGCGGAGGACAGACTGTTCAGGACCGCTCAGACGAGGGG 4042
Db 4006 TGGACCTGCGCTGCGGTGCGACGCGGAGGACAGACTGTTCAGGACCGCTCAGACGAGGGG 4065
QY 4043 ACTGTGAGCCCATCTGCTGCCCAACAGTTTCGCGTGTGCGAGCGGCGAGTGTCTCTCA 4102
Db 4066 ACTGTGAGCCCATCTGCTGCCCAACAGTTTCGCGTGTGCGAGCGG -CAGTGTGTCTCA 4124
QY 4103 TCAAAAGCAGTGGACTCTCTTCCCGACTGTATCGAGGGCTCCGACGAGCTCATGTGTG 4162
Db 4125 TCAAAAGCAGTGGACTCTCTTCCCGACTGTATCGAGGGCTCCGACGAGCTCATGTGTG 4184
QY 4163 AAATCACCAAGCCCTCTCAGACAGCCGCGCCACAGCAGTGCATCGGGCCGCTCA 4222
Db 4185 AAATCACCAAGCCCTCTCAGACAGCCGCGCCACAGCAGTGCATCGGGCCGCTCA 4244
QY 4223 TTGSCATCATCTCTCTCTCTTCTGATGGTGTGTCTATTTTGTGTGCGAGCGCGTGG 4282
Db 4245 TTGSCATCATCTCTCTCTCTTCTGATGGTGTGTCTATTTTGTGTGCGAGCGCGTGG 4304
QY 4283 TGTGCCAGCGCTATGCGGGGCCAAACGGGCCCTTCCCGCAGAGTATGTTCAGCGGGACCC 4342
Db 4305 TGTGCCAGCGCTATGCGGGGCCAAACGGGCCCTTCCCGCAGAGTATGTTCAGCGGGACCC 4364
QY 4343 CGCAGTGGCCCTCAATTTTCATAGCCCCCGGCGGTTCGCCAGCATGGCCCTTTCACAGGCA 4402
Db 4365 CGCAGTGGCCCTCAATTTTCATAGCCCCCGGCGGTTCGCCAGCATGGCCCTTTCACAGGCA 4424
QY 4403 TCGCATCGGAAAGTCCATGATGAGTCTCCGTGAGCTGTATGGGGGCGGGGGGTGTC 4462
Db 4425 TCGCATCGGAAAGTCCATGATGAGTCTCCGTGAGCTGTATGGGGGCGGGGGGTGTC 4484
QY 4463 CCCTGTACGACCGGAACACGATCAGAGGGGCTGTCCAGAGCTCGTCCAGACGAGG 4522
Db 4485 CCCTGTACGACCGGAACACGATCAGAGGGGCTGTCCAGAGCTCGTCCAGACGAGG 4544
QY 4523 CCAGCTGTACCCCGCATCTGAAACCGCGGCCCTCCCGGGCCACAGGACCCCTCCCTGT 4582
Db 4545 CCAGCTGTACCCCGCATCTGAAACCGCGGCCCTCCCGGGCCACAGGACCCCTCCCTGT 4604
QY 4583 ACAACATGGAATGTTTCTACTCTTTCAAAACATTTCCGGCCACTGCGAGACCGGTACAGGCCCT 4642
Db 4605 ACAACATGGAATGTTTCTACTCTTTCAAAACATTTCCGGCCACTGCGAGACCGGTACAGGCCCT 4664

Db 947 GGCAGCCTTTCTTCCACACTCGCTGTGAGGAGCAATGGCGGCTGCTCCCACTGTGCC 1006
Qy 983 TGCTGTCCCAAGCAGACCTTTCTACATGTCGCTGCCCTGCCAGCGGTGTGACGTGCGAGG 1042
Db 1007 TGCTGTCCCAAGCAGACCTTTCTACATGTCGCTGCCCTGCCAGCGGTGTGACGTGCGAGG 1066
Qy 1043 ACAACGGCAGGACGTGTAAAGCAGAGCGAGAGGTGTGCTGTGCTGCCCGCGCGAGCGG 1102
Db 1067 ACNACGGCAGGACGTGTAAAGCAGAGCGAGAGGTGTGCTGTGCTGCCCGCGCGAGCGG 1126
Qy 1103 ACCTACGAGGATCTCGCTGTGACACCGCGGACTTCACCGACATCGTGTGCTGCGAGGTGAGC 1162
Db 1127 ACCTACGAGGATCTCGCTGTGAGACACCGCGGACTTTACCGACATCGTGTGCGAGGTGAGC 1186
Qy 1163 ACATCCGCGACCCCATTTGCCATCGACTAGGACCCCGCTAGAGGGCTATGTCTACTGAGACAG 1222
Db 1187 ACATCCGCGACCCCATTTGCCATCGACTAGGACCCCGCTAGAGGGCTATGTCTACTGAGACAG 1246
Qy 1223 ATGACAGGTGGCGCCCATCCGACGGGTACTCTGACCGGCTCTGCGGCGCAGACGCTGG 1282
Db 1247 ATGACAGGTGGCGCCCATCCGACGGGTACTCTGACCGGCTCTGCGGCGCAGACGCTGG 1306
Qy 1283 TCAACACCEAGATCAACGACCCCGATGCGATGCGGTGCGACTGGGTGGCGCGAACCCTCT 1342
Db 1307 TCAACACCEAGATCAACGACCCCGATGCGATGCGGTGCGACTGGGTGGCGCGAACCCTCT 1366
Qy 1343 ACTGACCGACACCGGCGACCGCCATCGAGTGAACCGGCTCAACCGGACCTCCCGCA 1402
Db 1367 ACTGACCGACACCGGCGACCGCCATCGAGTGAACCGGCTCAACCGGACCTCCCGCA 1426
Qy 1403 AGATCCTGTGTGCGAGACCTGGACGACCGCGGACCATCGACTGCAACCGCTGTATGG 1462
Db 1427 AGATCCTGTGTGCGAGACCTGGACGACCGCGGACCATCGACTGCAACCGCTGTATGG 1486
Qy 1463 GCCTCATGTACTTGGACAGACTGGGAGAGAACCTTAAATCGAGTGTGCCAATTTGGATG 1522
Db 1487 GCCTCATGTACTTGGACAGACTGGGAGAGAACCTTAAATCGAGTGTGCCAATTTGGATG 1546
Qy 1523 GGCAGAGCGGCTGTGCTGCTCAATGCTCCTCGGTGGCCCAACCGGCTGGCCCTGG 1582
Db 1547 GGCAGAGCGGCTGTGCTGCTCAATGCTCCTCGGTGGCCCAACCGGCTGGCCCTGG 1606
Qy 1583 ACCTCAGAGGGGAGCTTACTGGGAGAGCGCCAGACAGAACAGATCGAGGTGATCA 1642
Db 1607 ACCTCAGAGGGGAGCTTACTGGGAGAGCGCCAGACAGAACAGATCGAGGTGATCA 1666
Qy 1643 ATGTTGATGGACGAGAGCGGACCTTCTGGAGGACAAAGTCCCGCACATTTTCGGGT 1702
Db 1667 ATGTTGATGGACGAGAGCGGACCTTCTGGAGGACAAAGTCCCGCACATTTTCGGGT 1726
Qy 1703 TCACGCTGTGGGAGCTTCACTACTGACTGCTGCGGCGCGGCGGCGGATCGAGCGGG 1762
Db 1727 TCACGCTGTGGGAGCTTCACTACTGACTGCTGCGGCGCGGCGGCGGATCGAGCGGG 1786
Qy 1763 TGCACAAAGGTCAAGCCAGCGGAGACGTTCATATTGACAGCTGCGCGACCTGTATGGGGC 1822
Db 1787 TGCACAAAGGTCAAGCCAGCGGAGACGTTCATATTGACAGCTGCGCGACCTGTATGGGGC 1846
Qy 1823 TCAAAGCTGTGAATGTGCCAAGGTGTCGGAACCAACCCGTGTGCGGACAGGAACGGGG 1882
Db 1847 TCAAAGCTGTGAATGTGCCAAGGTGTCGGAACCAACCCGTGTGCGGACAGGAACGGGG 1906
Qy 1883 GGTGACGACCTGTGCTTCTTCAACCCAGCGAACCCGCTGTGGCTGCCCATCGGCC 1942
Db 1907 GGTGACGACCTGTGCTTCTTCAACCCAGCGAACCCGCTGTGGCTGCCCATCGGCC 1966
Qy 1943 TGGAGCTGCTGAGTGACATGAAGACCTGCATCTGCTGTGAGGCTTCTTGTCTTCAACA 2002
Db 1967 TGGAGCTGCTGAGTGACATGAAGACCTGCATCTGCTGTGAGGCTTCTTGTCTTCAACA 2026
Qy 2003 GCAGAGCGCCATCCACAGGATCTCCCTCGAGACCAATTAACAAAGAGTGGCGCATCCCGC 2062

Db 2027 GCAGAGCCGCCATCCACAGGAT-TCCCTCGAGACCAATAACAAACGACGTGGCCATCCCGC 2085
Qy 2063 TCACGGGCGTCAAGAGGCTCAGCCCTGAGACTTTGATGTGTCTCAACCAACATCTACT 2122
Db 2086 TCACGGGCGTCAAGAGGCTCAGCCCTGAGACTTTGATGTGTCTCAACCAACATCTACT 2145
Qy 2123 GGACAGAGCTCAGCTGAGGACCATCAGCCGCTTCATGAACGGGAGCTCGGTGGAGC 2182
Db 2146 GGACAGAGCTCAGCTGAGGACCATCAGCCGCTTCATGAACGGGAGCTCGGTGGAGC 2205
Qy 2183 ACCTGTGAGGTTTGGCTTGTGACTACCCCGAGGGCATGCCGTTGACTTGGATGGGCAAGA 2242
Db 2206 ACCTGTGAGGTTTGGCTTGTGACTACCCCGAGGGCATGCCGTTGACTTGGATGGGCAAGA 2265
Qy 2243 ACCTTACTTGGGCGGACACTCGGACCAACAGAAATCGAGTGGCGGCTGAGCGGGCAGT 2302
Db 2266 ACCTTACTTGGGCGGACACTCGGACCAACAGAAATCGAGTGGCGGCTGAGCGGGCAGT 2325
Qy 2303 TCCGCAAGTCTCTGTGTGAGGAGCTTTGGACAAACCCGAGGTTCGCTGGCCCTGGATCCCA 2362
Db 2326 TCCGCAAGTCTCTGTGTGAGGAGCTTTGGACAAACCCGAGGTTCGCTGGCCCTGGATCCCA 2385
Qy 2363 CCAAGGGCTACATCTACTTGGACCCGAGTGGGCGGCAACCGAGGATCGTGGGGCTTCA 2422
Db 2386 CCAAGGGCTACATCTACTTGGACCCGAGTGGGCGGCAACCGAGGATCGTGGGGCTTCA 2445
Qy 2423 TGGACGGGACCAACTGCAATGACGCTGTGGAACAGTGGGCGGCGCAACGACCTCACA 2482
Db 2446 TGGACGGGACCAACTGCAATGACGCTGTGGAACAGTGGGCGGCGCAACGACCTCACA 2505
Qy 2483 TTGACTACGCTGACACGCGCTTACTTGGACCCGACTTGGACACCAACATGATCGAGTCTG 2542
Db 2506 TTGACTACGCTGACACGCGCTTACTTGGACCCGACTTGGACACCAACATGATCGAGTCTG 2565
Qy 2543 CCAACATGCTGGGTCAAGAGCGGCTGTGATTTGCGAGCATCTCCCGCACCCGTTTCGCTC 2602
Db 2566 CCAACATGCTGGGTCAAGAGCGGCTGTGATTTGCGAGCATCTCCCGCACCCGTTTCGCTC 2625
Qy 2603 TGAACGAGTACAGGATTTATATCTACTTGGACAGACTGGAATCTGACAGCATGAGCGGG 2662
Db 2626 TGAACGAGTACAGGATTTATATCTACTTGGACAGACTGGAATCTGACAGCATGAGCGGG 2685
Qy 2663 CCGACAGACTAGCGGCGGCAACCGCACCTTATCCAGGCGCACCTGACTTTCGTGATGG 2722
Db 2686 CCGACAGACTAGCGGCGGCAACCGCACCTTATCCAGGCGCACCTGACTTTCGTGATGG 2745
Qy 2723 ACATCTGGTGTTCATCTCTCCCGCAGAGTGGCTCAATGACTGTATGACACAAACG 2782
Db 2746 ACATCTGGTGTTCATCTCTCCCGCAGAGTGGCTCAATGACTGTATGACACAAACG 2805
Qy 2783 GGCAGTGTGGCAGCTGTGCTTGCATCCCGGGGCGCACCGCTGGGCTGGCGCTCAC 2842
Db 2806 GGCAGTGTGGCAGCTGTGCTTGCATCCCGGGGCGCACCGCTGGGCTGGCGCTCAC 2865
Qy 2843 ACTACACCTCGGACCCAGCAGCGCGCAACTGACGCGCGCCCAACCACTTTCGTGTTCA 2902
Db 2866 ACTACACCTCGGACCCAGCAGCGCGCAACTGACGCGCGCCCAACCACTTTCGTGTTCA 2925
Qy 2903 GCCAGAAATCTGCCATGATCGGATGATCCCGAGCAGACAGACCGCGGATCTATCC 2962
Db 2926 GCCAGAAATCTGCCATGATCGGATGATCCCGAGCAGACAGACCGCGGATCTATCC 2985
Qy 2963 TGCCCTTGCATGAGACTGAGGAACGTCAAAGCCATCGACTATGACCCACTGGACAGTTCA 3022
Db 2986 TGCCCTTGCATGAGACTGAGGAACGTCAAAGCCATCGACTATGACCCACTGGACAGTTCA 3045
Qy 3023 TCTACTGGGTGGATGGCGCCAGAACATCAAGCAGGCGCAAGCAGCGGACCCAGCCCT 3082
Db 3046 TCTACTGGGTGGATGGCGCCAGAACATCAAGCAGGCGCAAGCAGCGGACCCAGCCCT 3105
Qy 3083 TTGTTTGAACCTCTCTGAGCGCAAGCCAAACCCAGACAGGCGCCCAACGACCTCAGCA 3142
Db 3106 TTGTTTGAACCTCTCTGAGCGCAAGCCAAACCCAGACAGGCGCCCAACGACCTCAGCA 3165

APPLICANT: Kawaguchi, Yoshihiko
APPLICANT: Merriman, Tony R
APPLICANT: Metzker, Michael L
TITLE OF INVENTION: No. 6545137el Receptor
NUMBER OF SEQUENCES: 455
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon and Vanderhye
STREET: 1100 No. 6545137th Giebe Road, Eighth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: US
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/060,299
FILING DATE: 15-APR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/043,553
FILING DATE: 15-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/048,740
FILING DATE: 05-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: B.J. Sadoff
REGISTRATION NUMBER: 36,653
REFERENCE/DOCKET NUMBER: 620-35
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4091
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 5166 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-060-299-23

Query Match 96.6%; Score 4890; DB 4; Length 5166;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4915; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

Qy	142	GCCTCGCGCTCTGCTATTATTTGCCAAACCGCCGGGACGTACGGCTGGTGGAGACCGCGCGGA	201
Db	217	GCCTCGCGCTCTGCTATTATTTGCCAAACCGCCGGGACGTACGGCTGGTGGAGACCGCGCGGA	276
Qy	202	GTCAGCTGGAGTCCACATCGTGGTCAAGCGGCTGGAGGATGCGCCGCGAGTGACATTC	261
Db	277	GTCAGCTGGAGTCCACATCGTGGTCAAGCGGCTGGAGGATGCGCCGCGAGTGACATTC	336
Qy	262	CAGTTTTCGAAGGAGCGGTGTACTGGACAGACGTGACGAGGAGGCCATCAAGCAGACC	321
Db	337	CAGTTTTCGAAGGAGCGGTGTACTGGACAGACGTGACGAGGAGGCCATCAAGCAGACC	396
Qy	322	TACTGTAACAGACGGGGCGCGTGCAGAACTGGTCACTCCGGCTGCTCTCTCCC	381
Db	397	TACTGTAACAGACGGGGCGCGTGCAGAACTGGTCACTCCGGCTGCTCTCTCCC	456
Qy	382	GACGGCTCGCTCGACTGGGTGGGCAAGAAGCTGTACTGGACGGACTCAGAGACCAAC	441
Db	457	GACGGCTCGCTCGACTGGGTGGGCAAGAAGCTGTACTGGACGGACTCAGAGACCAAC	516
Qy	442	CGCATCAGAGTGGCAACCTCAATGGGCACATCCGGAAGGTGCTCTTTCTGGCAGAACCTT	501
Db	517	CGCATCAGAGTGGCAACCTCAATGGGCACATCCGGAAGGTGCTCTTTCTGGCAGAACCTT	576
Qy	502	GACCAGCTAGGGCCATCGCTTGGACCCCGCTCACGGGTACATGTACTGACACAGACTGG	561
Db	577	GACCAGCTAGGGCCATCGCTTGGACCCCGCTCACGGGTACATGTACTGACACAGACTGG	636

Db	1717	AA	TGTTGATGGGACGAAGGCGGACCTCTCTGGAGGCAAGCTCCCGCAATTTTCGGG	1776
Qy	1702	TT	CACGCTGCTGGGGGACTTCATCTACTGGACTGACTGGCAGCGCCGACGATCAGCGG	1761
Db	1777	TT	CACGCTGCTGGGGGACTTCATCTACTGGACTGACTGGCAGCGCCGACGATCAGCGG	1836
Qy	1762	GT	GCAACAAGGTCAAGGCCAGCGCGGAGGTGATCATTTGACACAGCTGCCGACCTGATGGG	1821
Db	1837	GT	GCAACAAGGTCAAGGCCAGCGCGGAGGTGATCATTTGACACAGCTGCCGACCTGATGGG	1896
Qy	1822	CT	CAAGCTGTGAATGTGGCCAAAGCGTCGTGGAACCAACCGCTGTGCGGACAGGAACGGG	1881
Db	1897	CT	CAAGCTGTGAATGTGGCCAAAGGTGTGGAACCAACCGCTGTGCGGACAGGAACGGG	1956
Qy	1882	GG	GTGACGCCACCTGTGTCTTTCACACCCACGCAACCCGGTGTGCTGCCCATCGGC	1941
Db	1957	GG	GTGACGCCACCTGTGTCTTTCACACCCACGCAACCCGGTGTGCTGCCCATCGGC	2016
Qy	1942	CT	GAGCTGCTGAGTGACATGAAGACCTGATCGTGTGCTGAGGCGCTTCTTGGTTTCAC	2001
Db	2017	CT	GAGCTGCTGAGTGACATGAAGACCTGATCGTGTGCTGAGGCGCTTCTTGGTTTCAC	2075
Qy	2002	AG	CAGCGCCCATCCACAGATCTCCCTCGAGACCAATACACAGCGTGGCCATCCCG	2061
Db	2076	AG	CAGCGCCCATCCACAGATCTCCCTCGAGACCAATACACAGCGTGGCCATCCCG	2135
Qy	2062	CT	CAGCGCGCTCAAGGAGGCTCAGCGCTTGACTTGTGTGTCCAAACACATCTAC	2121
Db	2136	CT	CAGCGCGCTCAAGGAGGCTCAGCGCTTGACTTGTGTGTCCAAACACATCTAC	2195
Qy	2122	TG	GACAGCTCAGCTTGAAAGACCATCAGCGCGCTTCATGAACCGGAGCTCGGTGGAG	2181
Db	2196	TG	GACAGCTCAGCTTGAAAGACCATCAGCGCGCTTCATGAACCGGAGCTCGGTGGAG	2255
Qy	2182	CA	GTGTGTGAGTTTGGCGCTTGACTACCCGAGGCGATGCGCGTTGACTGATGGGCAAG	2241
Db	2256	CA	GTGTGTGAGTTTGGCGCTTGACTACCCGAGGCGATGCGCGTTGACTGATGGGCAAG	2315
Qy	2242	AA	CTCTACTTGGGCGGACACTTGGGACCAAACAGAAATCGAAGTGGCGCGCTGGACGGGCG	2301
Db	2316	AA	CTCTACTTGGGCGGACACTTGGGACCAAACAGAAATCGAAGTGGCGCGCTGGACGGGCG	2375
Qy	2302	TT	CGGCAAGTCTCTGTGTGAGGAGCTTGGAACAACCCGAGTGCCTGGGCTCTGGATCCC	2361
Db	2376	TT	CGGCAAGTCTCTGTGTGAGGAGCTTGGAACAACCCGAGTGCCTGGGCTCTGGATCCC	2435
Qy	2362	AC	NAAGGCTACATCTACTGGAACGAGTGGGCGGGAAGCCGAGGATCGTGGCGGCTTC	2421
Db	2436	AC	NAAGGCTACATCTACTGGAACGAGTGGGCGGGAAGCCGAGGATCGTGGCGGCTTC	2495
Qy	2422	AT	GACGGGACCAACTGTCATGACGCTGTTGGAACAAGTGGCGGGGCAACGACCTCACC	2481
Db	2496	AT	GACGGGACCAACTGTCATGACGCTGTTGGAACAAGTGGCGGGGCAACGACCTCACC	2555
Qy	2482	ATT	GACTACGCTGACACGCGCTCTACTGGACCGACCTGGACACCAACATGATCGAGTCG	2541
Db	2556	ATT	GACTACGCTGACACGCGCTCTACTGGACCGACCTGGACACCAACATGATCGAGTCG	2615
Qy	2542	TC	CAATGCTGGGTCAAGGAGCGGTCGTGATTTGCCGACGATCTCCCGACCGGTTCCGGT	2601
Db	2616	TC	CAATGCTGGGTCAAGGAGCGGTCGTGATTTGCCGACGATCTCCCGACCGGTTCCGGT	2675
Qy	2602	CT	GACGATCAGCGATTAATCTACTGACACGACTGGAAATCTGCACACGATTTGACGG	2661
Db	2676	CT	GACGATCAGCGATTAATCTACTGACACGACTGGAAATCTGCACACGATTTGACGG	2735
Qy	2662	GC	CAACAAGTACGCGGCCGGAACCGCACCTCATCCAGGCGCACCTGGACTTCGTGATG	2721
Db	2736	GC	CAACAAGTACGCGGCCGGAACCGCACCTCATCCAGGCGCACCTGGACTTCGTGATG	2795
Qy	2722	GA	CATCTGTGTTCCTCTCTCCGCGCAGGATGGCTCAATGACTGATGTGCAACAAC	2781

2796	QY	GGGAGTGTGGG	AGCTGTGCTTTGCCATCCCGGGGGCCACCGCTGCGGCTCGGCTCA	2841
2856	QY	GGGAGTGTGGG	AGCTGTGCTTTGCCATCCCGGGGGCCACCGCTGCGGCTCGGCTCA	2915
2842	QY	CACATACACCTTGGAC	CCCGAGCGCAACTGCGAGCCGGCCACCACTTCTTGTCTGTTCT	2901
2916	QY	CACATACACCTTGGAC	CCCGAGCGCAACTGCGAGCCGGCCACCACTTCTTGTCTGTTCT	2975
2902	QY	AGCCAGAAATCTGCA	TCATCAGTCCGGACGACACGACGACGACGCCCGGATCTCATC	2961
2976	QY	AGCCAGAAATCTGCA	TCATCAGTCCGGACGACACGACGACGACGCCCGGATCTCATC	3035
2962	QY	CTGCCCTCTGATGGA	CTGAGAACGTCCTCAAGCCATCGACTATGACCCACTGGACAAGTTC	3021
3036	QY	CTGCCCTCTGATGGA	CTGAGAACGTCCTCAAGCCATCGACTATGACCCACTGGACAAGTTC	3095
3022	QY	ATCTACTGGGTGGAT	TGGGCGCGAGAAATCAAGGAGCGAAGGACGACGAGGCCAGCC	3081
3096	QY	ATCTACTGGGTGGAT	TGGGCGCGAGAAATCAAGGAGCGAAGGACGACGAGGCCAGCC	3155
3082	QY	TTTGTGTTTGACCT	CTCTGAGCCAAAGGCCAAAACCCAGACAGGACGACCCACGACTCAGC	3141
3156	QY	TTTGTGTTTGACCT	CTCTGAGCCAAAGGCCAAAACCCAGACAGGACGACCCACGACTCAGC	3215
3142	QY	ATCGACATCTACAG	CCCGGACACTGTTCTGGACGCTGGAGGCCACCAATACCATCAAGTTC	3201
3216	QY	ATCGACATCTACAG	CCCGGACACTGTTCTGGACGCTGGAGGCCACCAATACCATCAAGTTC	3275
3202	QY	CACAGGCTGAGCGGG	GAAGCATGGGGGTGGTGTGGGTGGGACCGCGACCAAGCCGAGG	3261
3276	QY	CACAGGCTGAGCGGG	GAAGCATGGGGGTGGTGTGGGTGGGACCGCGACCAAGCCGAGG	3335
3262	QY	GCCATCGTCTCAAC	CGCGGACGAGGAGTCTCTTCTTCAACCAATGACGAGGACCGGCA	3321
3336	QY	GCCATCGTCTCAAC	CGCGGACGAGGAGTCTCTTCTTCAACCAATGACGAGGACCGGCA	3395
3322	QY	GCCAAAGATCGAA	CGCGCAGCCCTGACCGGACCGGAGCGGAGGCTCTCTTCAACCGCGC	3381
3396	QY	GCCAAAGATCGAA	CGCGCAGCCCTGACCGGACCGGAGCGGAGGCTCTCTTCAACCGCGC	3455
3382	QY	CTCATCCGCCCTGT	GGGCCCTGGGTGGGACAAACACATGGGGCAAGCTGTTCTGGGTGGAC	3441
3456	QY	CTCATCCGCCCTGT	GGGCCCTGGGTGGGACAAACACATGGGGCAAGCTGTTCTGGGTGGAC	3515
3442	QY	CGCGACCTGAGCGC	ATTGAGCTGTGACCTCTCAGGGGCCAACCGCTGACCCCTGGAG	3501
3516	QY	CGCGACCTGAGCGC	ATTGAGCTGTGACCTCTCAGGGGCCAACCGCTGACCCCTGGAG	3575
3502	QY	GACGCCAAATCTGT	GACGCTCTGGGGCTGACCATCTTGGGCAAGATCTCTACTGGATC	3561
3576	QY	GACGCCAAATCTGT	GACGCTCTGGGGCTGACCATCTTGGGCAAGATCTCTACTGGATC	3635
3562	QY	GACCGCAGCAGCAG	ATGATCGAGCGGTGGGAAAGACCAACCGGGGACCAAGCGGACTCGC	3621
3636	QY	GACCGCAGCAGCAG	ATGATCGAGCGGTGGGAAAGACCAACCGGGGACCAAGCGGACTCGC	3695
3622	QY	ATCCAGGGCCGTGT	CGGCCCACTTCACTGGGCATCATGACGATGAGGAGGATGACGCTGGAG	3681
3696	QY	ATCCAGGGCCGTGT	CGGCCCACTTCACTGGGCATCATGACGATGAGGAGGATGACGCTGGAG	3755
3682	QY	GAGTTCTCAGCCAC	CCCATGTGCCGTGACAAATGGTGGCTGCTCCACATCTCTATTGTC	3741
3756	QY	GAGTTCTCAGCCAC	CCCATGTGCCGTGACAAATGGTGGCTGCTCCACATCTCTATTGTC	3815
3742	QY	AAGGGTGATGGGAC	CACCAACGCTGCTATGCCAGTCCACTCGTGTCTCTGACGAACTCTG	3801
3816	QY	AAGGGTGATGGGAC	CACCAACGCTGCTATGCCAGTCCACTCGTGTCTCTGACGAACTCTG	3875
3802	QY	CTGACTGTGGAGAG	CGCCGCCACTTGTCTCCCGGACCAAGTTTGCATGTGCGACAGGGGAG	3861
3876	QY	CTGACTGTGGAGAG	CGCCGCCACTTGTCTCCCGGACCAAGTTTGCATGTGCGACAGGGGAG	3935

Qy 3862 ATCGACTGTATCCCGGGGCTGGCGCTGTGACGGCTTTCCGAGTGGATGACACAGAGC 3921
Db |||||
Qy 3936 ATCGACTGTATCCCGGGGCTGGCGCTGTGACGGCTTTCCGAGTGGATGACACAGAGC 3995
Db |||||
Qy 3922 GACGAGGAGGCTGCCCGTGTGCTCCGCGCCAGTTCCCTGCGCGCGGGGTGAGTGT 3981
Db |||||
Qy 3996 GACGAGGAGGCTGCCCGTGTGCTCCGCGCCAGTTCCCTGCGCGCGGGGTGAGTGT 4055
Db |||||
Qy 3982 GTGGACCTGCGCGCTGCGACGCGCGAGGAGAGACTGTGAGGACCGCTCAGACGAGCGC 4041
Db |||||
Qy 4056 GTGGACCTGCGCGCTGCGACGCGCGAGGAGAGACTGTGAGGACCGCTCAGAC-AGGCG 4114
Db |||||
Qy 4042 GACTGTGACGCACTGCTGCTGCCCAACAGTTCGGTGTGCGAGCGCCAGTGTGTCCTC 4101
Db |||||
Qy 4115 GACTGTGACGCACTGCTGCTGCCCAACAGTTCGGTGTGCGAGCGCCAGTGTGTCCTC 4174
Db |||||
Qy 4102 ATCAACAGCAGTGGAGTCTCTTCCCGAGTGTATCGACGGCTCCGACGAGCTCATGTGT 4161
Db |||||
Qy 4175 ATCAACAGCAGTGGAGTCTCTTCCCGAGTGTATCGACGGCTCCGACGAGCTCATGTGT 4234
Db |||||
Qy 4162 GAAATCACCAAGCGCCCTCAGACGACAGCCCGGCCACAGCAGTGCATCGGGCCCGTC 4221
Db |||||
Qy 4235 GAAATCACCAAGCGCCCTCAGACGACAGCCCGGCCACAGCAGTGCATCGGGCCCGTC 4294
Db |||||
Qy 4222 ATTGGCATCATCTCTCTCTCTCTGTCATGGTGTGTCTATTTGTGTGTCAGCGCGTG 4281
Db |||||
Qy 4295 ATTGGCATCATCTCTCTCTCTGTCATGGTGTGTCTATTTGTGTGTCAGCGCGTG 4354
Db |||||
Qy 4282 GTGTGCCAGCGCTATGCGGGGGCCAAAGCGGCCCTTCCCGCAGCAGTATGTCAGCGGAGC 4341
Db |||||
Qy 4355 GTGTGCCAGCGCTATGCGGGGGCCAAAGCGGCCCTTCCCGCAGCAGTATGTCAGCGGAGC 4414
Db |||||
Qy 4342 CGCACGTGCCCTCAATTTTATAGCCCCCGGGCGGTTCACAGCATGCCCCCTTCACAGGC 4401
Db |||||
Qy 4415 CGCACGTGCCCTCAATTTTATAGCCCCCGGGCGGTTCACAGCATGCCCCCTTCACAGGC 4474
Db |||||
Qy 4402 ATCGCATCGGAAAGTCCATGATGAGTCCGTGAGCCTGATGGGGGGCGGGGGGGTG 4461
Db |||||
Qy 4475 ATCGCATCGGAAAGTCCATGATGAGTCCGTGAGCCTGATGGGGGGCGGGGGGGTG 4534
Db |||||
Qy 4462 CCCCTGTACGACCGGAACACAGTCCAGCGGGCTCGTCCAGCAGTCCGTCCAGCAGCAAG 4521
Db |||||
Qy 4535 CCCCTGTACGACCGGAACACAGTCCAGCGGGCTCGTCCAGCAGTCCGTCCAGCAGCAAG 4594
Db |||||
Qy 4522 GCCAGCTGTACCGCGCGATCTTGAAACCGCGCGCTCCCGGGCCACGAGCCCTCCCTG 4581
Db |||||
Qy 4595 GCCAGCTGTACCGCGCGATCTTGAAACCGCGCGCTCCCGGGCCACGAGCCCTCCCTG 4654
Db |||||
Qy 4582 TACAACATGGACATGTTCTACTCTTCAACATTCGGGCCACTGCGAGCCGTACAGGCC 4641
Db |||||
Qy 4655 TACAACATGGACATGTTCTACTCTTCAACATTCGGGCCACTGCGAGCCGTACAGGCC 4714
Db |||||
Qy 4642 TACATCATTCGAGGAATGGCGCCCGCCGACGAGCGCCCTGACAGCCAGCTGTGTGACAGC 4701
Db |||||
Qy 4715 TACATCATTCGAGGAATGGCGCCCGCCGACGAGCGCCCTGACAGCCAGCTGTGTGACAGC 4774
Db |||||
Qy 4702 GACTACAGCGCGAGCGCTGGAAGGCCACGAGTACTACTTGGATTTGAACTCGGACTCA 4761
Db |||||
Qy 4775 GACTACAGCGCGAGCGCTGGAAGGCCACGAGTACTACTTGGATTTGAACTCGGACTCA 4834
Db |||||
Qy 4762 GACCCCTATCCACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCA 4821
Db |||||
Qy 4835 GACCCCTATCCACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCA 4894
Db |||||
Qy 4822 CCGCCCTCGCCCGCCACCGAGGAGTACTTCCATCTTCCCGCCCTCCCTCCCTCCCTCC 4881
Db |||||
Qy 4895 CCGCCCTCGCCCGCCACCGAGGAGTACTTCCATCTTCCCGCCCTCCCTCCCTCCCTCC 4954
Db |||||
Qy 4882 TGCAGGAGTCACTGACCTCGCGCGGGCCACTCTGCTGCTTCTGTCGCCCTGTAATA 4941
Db |||||
Qy 4955 TGCAGGAGTCACTGACCTCGCGCGGGCCACTCTGCTGCTTCTGTCGCCCTGTAATA 5014
Db |||||

Qy 4942 GTTTTAAATATGAACAAAGAAAAATATATTTTATGATTTAAAAAATAAATATATTTGG 5001
Db |||||
Qy 5015 GTTTTAAATATGAACAAAGAAAAATATATTTTATGATTTAAAAAATAAATATATTTGG 5074
Db |||||
Qy 5002 GATTTTAAAAACATGAGAAATGTGAACCTGTGATGGGTGGCGAGGCTGGGAGAACTTTG 5061
Db |||||
Qy 5075 GATTTTAAAAACATGAGAAATGTGAACCTGTGATGGGTGGCGAGGCTGGGAGAACTTTG 5134
Db |||||
Qy 5062 TA 5063
Db ||
Qy 5135 TA 5136
Db ||

RESULT 15

US-09-402-923A-23
; Sequence 23, Application US/09402923A
; Patent No. 6555654
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; Hess, John W
; Caskey, Charles T
; Cox, Roger D
; Gerhold, David
; Hammond, Holly
; Hey, Patricia
; Kawaguchi, Yoshihiko
; Merriman, Tony R
; Metzker, Michael L
; TITLE OF INVENTION: No. 6555654el LDL-Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6555654th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; ZIP: VA 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/402,923A
; FILING DATE: 14-Feb-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB98/01102
; FILING DATE: 15-APR-1998
; APPLICATION NUMBER: US 60/043,553
; FILING DATE: 15-APR-1997
; APPLICATION NUMBER: US 60/048,740
; FILING DATE: 05-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: B. J. Sadoff
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 620-81
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4091
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5166 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-402-923A-23

Query Match 96.6%; Score 4890; DB 4; Length 5166;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4915; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

Qy 142 GCCTCGCGCTCCTGCTATTGTCACCGCGGACGTACGGCTGGTGACCGCGGCGGA 201

Db 217 GCCTCGCGCTCCTGCTATTTGCAAAACCGCCGGGACGTACGGCTGGTGGACGCCGGCGGA 276
QY 202 GTCAGCTGGAGTCCACCATCGTGGTCAGCGGCCCTGGAGGATCGGCCCGCAGTGGACTTC 261
Db 277 GTCAGCTGGAGTCCACCATCGTGGTCAGCGGCCCTGGAGGATCGGCCCGCAGTGGACTTC 336
QY 262 CAGTTTTTCCAAAGGAGCGCTGTACTGGACAGACGTGACGAGGAGGCCCATCAAGCAGAC 321
Db 337 CAGTTTTTCCAAAGGAGCGCTGTACTGGACAGACGTGACGAGGAGGCCCATCAAGCAGAC 396
QY 322 TACCTGAACCAAGACGGGGCCCGCTGTGACAGAACTGTGATCTCGGGCTCGTCTCTCCC 381
Db 397 TACCTGAACCAAGACGGGGCCCGCTGTGACAGAACTGTGATCTCGGGCTCGTCTCTCCC 456
QY 382 GACGSCCTCGCTCGGACTGGGTGGGCAAGAGCTGTACTGGACGGAATCAAGACCAAC 441
Db 457 GACGSCCTCGCTCGGACTGGGTGGGCAAGAGCTGTACTGGACGGAATCAAGACCAAC 516
QY 442 CGCATCGAGGTGGCCAACTCAATGGCACATCCCGAAGGTGCTTCTTGGCAGGACCTT 501
Db 517 CGCATCGAGGTGGCCAACTCAATGGCACATCCCGAAGGTGCTTCTTGGCAGGACCTT 576
QY 502 GACCAAGCTAGGGCCATCGCTTGGACCCCGCTCAACGGGTACATGTACTGGACAGACTGG 561
Db 577 GACCAAGCTAGGGCCATCGCTTGGACCCCGCTCAACGGGTACATGTACTGGACAGACTGG 636
QY 562 GGTGAGACCCCGGATTTGAGGGGCGAGGATGATGACAGCACCCCGAAGATCATTTGTG 621
Db 637 GGTGAGACCCCGGATTTGAGGGGCGAGGATGATGACAGCACCCCGAAGATCATTTGTG 696
QY 622 GACTCGGACATTTACTGGCCCAATGGACTGACATCGACCTGGAGGACAGAGCTCTAC 681
Db 697 GACTCGGACATTTACTGGCCCAATGGACTGACATCGACCTGGAGGACAGAGCTCTAC 756
QY 682 TGGGCTGACGCAAGCTCAGCTTTCATCCACCGCTGCCAACTTGGACGGCTCGTTCCGGCAG 741
Db 757 TGGGCTGACGCAAGCTCAGCTTTCATCCACCGCTGCCAACTTGGACGGCTCGTTCCGGCAG 816
QY 742 AAGGTGGTGGAGGGGAGCTGACGACCCCTTGGCCCTTGACGCTTCCGGGACACTCTG 801
Db 817 AAGGTGGTGGAGGGGAGCTGACGACCCCTTGGCCCTTGACGCTTCCGGGACACTCTG 876
QY 802 TACTGGACAGACTGSGACACCGCTTCATTCATGCTGCAACAGCGCACTGSGGGGAG 861
Db 877 TACTGGACAGACTGSGACACCGCTTCATTCATGCTGCAACAGCGCACTGSGGGGAG 936
QY 862 AGGAAGGAGATCCTGAGTGCCTCTTACTCACCCATGGACATCCAGGTGCTGAGCCAGGAG 921
Db 937 AGGAAGGAGATCCTGAGTGCCTCTTACTCACCCATGGACATCCAGGTGCTGAGCCAGGAG 996
QY 922 CGGACGCTTCTTCCACACTGCTGTGAGGAGCAATGSGGCTGCTCCACCTGTGC 981
Db 997 CGGACGCTTCTTCCACACTGCTGTGAGGAGCAATGSGGCTGCTCCACCTGTGC 1056
QY 982 CTGCTGTCCCAAGCGAGCTTTCTACACATGCGCTGCCCCACGGGTGCGAGTGCAG 1041
Db 1057 CTGCTGTCCCAAGCGAGCTTTCTACACATGCGCTGCCCCACGGGTGCGAGTGCAG 1116
QY 1042 GACAAACGACGAGCGTGTAAAGCAGGAGCCGAGGAGGTGCTGCTGCTGGCCCGCGGACG 1101
Db 1117 GACAAACGACGAGCGTGTAAAGCAGGAGCCGAGGAGGTGCTGCTGCTGGCCCGCGGAGC 1176
QY 1102 GACCTACGAGGATCTCGCTGGAACACGCGGACTTACCGGACATGCTGCTGAGGTGGAC 1161
Db 1177 GACCTACGAGGATCTCGCTGGAACACGCGGACTTACCGGACATGCTGCTGAGGTGGAC 1236
QY 1162 GACATCGGACGCAATTCATTCGACTACGACCGCTAGAGGGCTATGCTACTTGGACA 1221
Db 1237 GACATCGGACGCAATTCGACTACGACCGCTAGAGGGCTATGCTACTTGGACA 1296
QY 1222 GATGACAGAGGTGCGGGCCATCCGACAGGGCGTACTTGGACGGGTCTGGGGCGCAGACGCTG 1281

Db 1297 GATGACAGAGGTGCGGGCCATCCGACAGGGCGTACTCTGGACGGGTCTTGGGGCGCAGACGCTG 1356
QY 1282 GTCAAACCCGAGATCAACGACCCCGATGGCATCGGGTCGACTGGGTGGCCCGGAACCTC 1341
Db 1357 GTCAAACCCGAGATCAACGACCCCGATGGCATCGGGTCGACTGGGTGGCCCGGAACCTC 1416
QY 1342 TACTGGACCCGACACGCGGCACCGCATCGAGGTGACGCGCCTCAACGGCACCTTCCGCG 1401
Db 1417 TACTGGACCCGACACGCGGCACCGCATCGAGGTGACGCGCCTCAACGGCACCTTCCGCG 1476
QY 1402 AAGATCTCGTGTGTCGAGGACCTTGGACGAGCCCGAGCCATCGCACTGCAACCCCGTATG 1461
Db 1477 AAGATCTCGTGTGTCGAGGACCTTGGACGAGCCCGAGCCATCGCACTGCAACCCCGTATG 1536
QY 1462 GGCCTCATGTACTGCAAGACTGGGAGAGAACCTTAAATCGAGTGTGCCAATTTGAT 1521
Db 1537 GGCCTCATGTACTGCAAGACTGGGAGAGAACCTTAAATCGAGTGTGCCAATTTGAT 1596
QY 1522 GGGCAGGAGCGGCGTGTGCTGGTCAATGCTCTTCCCTCGGGTGGCCCAACGGCCTGGCCCTG 1581
Db 1597 GGGCAGGAGCGGCGTGTGCTGGTCAATGCTCTTCCCTCGGGTGGCCCAACGGCCTGGCCCTG 1656
QY 1582 GACCTGAGGAGGGGAAAGCTCTACTGGGAGACGCGCAAGACAGACATCGAGGTATC 1641
Db 1657 GACCTGAGGAGGGGAAAGCTCTACTGGGAGACGCGCAAGACAGACATCGAGGTATC 1716
QY 1642 AATGTTGATGGGACGAGAGCGGACCTCTCTGGAGGACAAAGCTCCCGCACATTTTCGGG 1701
Db 1717 AATGTTGATGGGACGAGAGCGGACCTCTCTGGAGGACAAAGCTCCCGCACATTTTCGGG 1776
QY 1702 TTCACTGCTGTGGGGGACTTTCATCTACTGGACTGACTGGCAGCGCCGACGATCGAGGG 1761
Db 1777 TTCACTGCTGTGGGGGACTTTCATCTACTGGACTGACTGGCAGCGCCGACGATCGAGGG 1836
QY 1762 GTGCAAGAGTCAAGGCGACGCGGAGCGTCAATCAATGACAGCTGCGCCGACTGATGGGG 1821
Db 1837 GTGCAAGAGTCAAGGCGACGCGGAGCGTCAATCAATGACAGCTGCGCCGACTGATGGGG 1896
QY 1822 CTCAAAGCTGTAATGTGGCCAAAGTGTGCGGAACCAACCGCTGTGCGGACAGGAACGGG 1881
Db 1897 CTCAAAGCTGTAATGTGGCCAAAGTGTGCGGAACCAACCGCTGTGCGGACAGGAACGGG 1956
QY 1882 GGGTGCAGCCACCTGTGCTTTCACACCCACGCAACCCCGTGTGGCTGCGCCCATCGGC 1941
Db 1957 GGGTGCAGCCACCTGTGCTTTCACACCCACGCAACCCCGTGTGGCTGCGCCCATCGGC 2016
QY 1942 CTGAGCTGTGATGACATGAAGACCTGCAATCGTGTGCTGAGGCGCTTCTTGGTCTTAC 2001
Db 2017 CTGAGCTGTGATGACATGAAGACCTGCAATCGTGTGAGGCGCTTCTTGGTCTTAC 2075
QY 2002 AGCAGCGCGCATCCACAGGATCTCCCTCGAGACCAATACAGAGCTGSCCATCCG 2061
Db 2076 AGCAGCGCGCATCCACAGGATCTCCCTCGAGACCAATACAGAGCTGSCCATCCG 2135
QY 2062 CTCACGGCGCTCAAGAGGCGCTCAGCCCTGCACTTTGATGTGTCCAAACACCATCTAC 2121
Db 2136 CTCACGGCGCTCAAGAGGCGCTCAGCCCTGCACTTTGATGTGTCCAAACACCATCTAC 2195
QY 2122 TGGACAGAGTCAAGCTGGAAGACCATCAGCGCGCTTTCATGAACGGGAGTCTGGTGGAG 2181
Db 2196 TGGACAGAGTCAAGCTGGAAGACCATCAGCGCGCTTTCATGAACGGGAGTCTGGTGGAG 2255
QY 2182 CACGTGTGGATTTGGCTTTGACTACCCCGAGGGCATGGCCGTGACTGGATGGGCAAG 2241
Db 2256 CACGTGTGGATTTGGCTTTGACTACCCCGAGGGCATGGCCGTGACTGGATGGGCAAG 2315
QY 2242 AACCTCTACTGGGCGCACATGGGACCAACAGAAATCGAAGTGGCGCGCTCGACCGGCGAG 2301
Db 2316 AACCTCTACTGGGCGCACATGGGACCAACAGAAATCGAAGTGGCGCGCTCGACCGGCGAG 2375
QY 2302 TTCCGGCAGAGTCTCGTGTGAGGAGACTTGAACAAACCGAGGTGCTGGCGCTGGATCCC 2361
Db 2376 TTCCGGCAGAGTCTCGTGTGAGGAGACTTGAACAAACCGAGGTGCTGGCGCTGGATCCC 2435

QY	2362	ACCAAGGGCTACATCTACTGGACCGAGTGGGGCGGCAAGCCGAGGATCGTGGGGCCCTTC	2421
Db	2436	ACCAAGGGCTACATCTACTGGACCGAGTGGGGCGGCAAGCCGAGGATCGTGGGGCCCTTC	2495
QY	2422	ATGGACGGGACCAACTGTCATGACGCTGTGGACAAGGTGGGCGGGCCAAAGCACTCACC	2481
Db	2496	ATGGACGGGACCAACTGTCATGACGCTGTGGACAAGGTGGGCGGGCCAAAGCACTCACC	2555
QY	2482	ATTGACTACGCTGACCGGCTCTACTGGACCGGACCTGGGACCAACATGATGAGTGG	2541
Db	2556	ATTGACTACGCTGACCGGCTCTACTGGACCGGACCTGGGACCAACATGATGAGTGG	2615
QY	2542	TCCACATGCTGGGTGACGAGGGGTCTGTGATGGGACGATCTCCCGACCCGCTTCGGT	2601
Db	2616	TCCACATGCTGGGTGACGAGGGGTCTGTGATGGGACGATCTCCCGACCCGCTTCGGT	2675
QY	2602	CTGACGCAAGTACAGCGATTATCTACTGGGACAGACTGGAATCTGCACAGCATTTGAGCGG	2661
Db	2676	CTGACGCAAGTACAGCGATTATCTACTGGGACAGACTGGAATCTGCACAGCATTTGAGCGG	2735
QY	2662	GCCGACAAGACTAGCGGCGGAAACCGCACCTCTATCCAGGGCCACCTGGACTTCGTGATG	2721
Db	2736	GCCGACAAGACTAGCGGCGGAAACCGCACCTCTATCCAGGGCCACCTGGACTTCGTGATG	2795
QY	2722	GACATCTGCTGTTCCACTCTCCCGCCAGGATGGCCCTCAATGACTGTATGCACACAAAC	2781
Db	2796	GACATCTGCTGTTCCACTCTCCCGCCAGGATGGCCCTCAATGACTGTATGCACACAAAC	2855
QY	2782	GGGCAAGTGTGGGCACTGTGCTTGGCCATCCCGGCGGCCACCGCTGGCGGTGGCGCTCA	2841
Db	2856	GGGCAAGTGTGGGCACTGTGCTTGGCCATCCCGGCGGCCACCGCTGGCGGTGGCGCTCA	2915
QY	2842	CACATACCCCTGGACCCAGCAGCGGCAACTGACGCCGCCCAACCACTTCTTGCTGTTTC	2901
Db	2916	CACATACCCCTGGACCCAGCAGCGGCAACTGACGCCGCCCAACCACTTCTTGCTGTTTC	2975
QY	2902	AGCCAGAAATCTGGCATCAGTTCGGATGATCCCGGACGACGACGACGACGACGACGACG	2961
Db	2976	AGCCAGAAATCTGGCATCAGTTCGGATGATCCCGGACGACGACGACGACGACGACGACG	3035
QY	2962	CTGCCCCCTGCATGGAAGTCTGAGGAACTGCAAGCCATCGACTATGACCCACTGGAAGTTC	3021
Db	3036	CTGCCCCCTGCATGGAAGTCTGAGGAACTGCAAGCCATCGACTATGACCCACTGGAAGTTC	3095
QY	3022	ATCTACTGGGTGATGGGCGCCAGAAATCAAGGAGGACCAAGGACGAGGACCCAGGCC	3081
Db	3096	ATCTACTGGGTGATGGGCGCCAGAAATCAAGGAGGACCAAGGACGAGGACCCAGGCC	3155
QY	3082	TTTGTGTTTGTGACTCTCTGAGCCAGGCGCAAAACCCAGACGAGGACGCCCAAGCTCAGC	3141
Db	3156	TTTGTGTTTGTGACTCTCTGAGCCAGGCGCAAAACCCAGACGAGGACGCCCAAGCTCAGC	3215
QY	3142	ATCGACATCTACGCGGACACTGTTCTGGACGTCGAGGCGCACCAATACCATCAACGTC	3201
Db	3216	ATCGACATCTACGCGGACACTGTTCTGGACGTCGAGGCGCACCAATACCATCAACGTC	3275
QY	3202	CACAGGCTGAGCGGGGAAGCAATGGGGGTGTGTGCTGGGACCGCGAACGCGCCAGG	3261
Db	3276	CACAGGCTGAGCGGGGAAGCAATGGGGGTGTGTGCTGGGACCGCGAACGCGCCAGG	3335
QY	3262	GCCATCGTCTGTCAGCGGAGGGGTACCTGTACTTCCACCAATGACGAGCGGGCA	3321
Db	3336	GCCATCGTCTGTCAGCGGAGGGGTACCTGTACTTCCACCAATGACGAGCGGGCA	3395
QY	3322	GCCAAAGTTCGAACGCGGACCTTGGACCGGACCCGAGCGGAGGTCTCTTCAACACCGGC	3381
Db	3396	GCCAAAGTTCGAACGCGGACCTTGGACCGGACCCGAGCGGAGGTCTCTTCAACACCGGC	3455
QY	3382	CTCATTCGCGCCCTGTGGCCCTGGTGTGGACAACACACTGGGGCAAGCTGTTCTGGGTGAC	3441
Db	3456	CTCATTCGCGCCCTGTGGCCCTGGTGTGGACAACACACTGGGGCAAGCTGTTCTGGGTGAC	3515

QY	3442	GCGGACCTGAGCGCATTTGAGAGCTGTGACCTGTGAGGGGCCAACCCTGACCTGACCTGGAG	3501
Db	3516	GCGGACCTGAGCGCATTTGAGAGCTGTGACCTGTGAGGGGCCAACCCTGACCTGACCTGGAG	3575
QY	3502	GACGCCAAACATCTGTGACGCTCTGGGCTGTGACCATCTTGGCAAGCATCTCTACTGGATC	3561
Db	3576	GACGCCAAACATCTGTGACGCTCTGGGCTGTGACCATCTTGGCAAGCATCTCTACTGGATC	3635
QY	3562	GACGCCAAGCAGCAGATGATCGAGGCTGTGAGAGAACCAACGGGGGAACAGCGGACTGCG	3621
Db	3636	GACGCCAAGCAGCAGATGATCGAGGCTGTGAGAGAACCAACGGGGGAACAGCGGACTGCG	3695
QY	3622	ATCCAGGGCCGTGTGCGCCACTCACTGCGCATCCATGCAAGTGGAGGAGTCAAGCTGAG	3681
Db	3696	ATCCAGGGCCGTGTGCGCCACTCACTGCGCATCCATGCAAGTGGAGGAGTCAAGCTGAG	3755
QY	3682	GAGTTCTCAGCCCAACCCATGTGCGCTGACAAATGGTGGCTGCTCCCACTCTGTATTGCC	3741
Db	3756	GAGTTCTCAGCCCAACCCATGTGCGCTGACAAATGGTGGCTGCTCCCACTCTGTATTGCC	3815
QY	3742	AAGGGTATGGGGAACACACGCTGTCTATGCCAGTCCACCTCTGCTGCTCTGCGAGAACCTG	3801
Db	3816	AAGGGTATGGGGAACACACGCTGTCTATGCCAGTCCACCTCTGCTGCTCTGCGAGAACCTG	3875
QY	3802	CTGACCTGTGGAGAGCGCCCACTGCTCCCGGACGAGTTTGCATGTGACACAGGGGAG	3861
Db	3876	CTGACCTGTGGAGAGCGCCCACTGCTCCCGGACGAGTTTGCATGTGACACAGGGGAG	3935
QY	3862	ATCGACTGTATCCCGGGGGCTGGCGCTGTGACGGCTTTCGCCAGTGCATGTGACACAGG	3921
Db	3936	ATCGACTGTATCCCGGGGGCTGGCGCTGTGACGGCTTTCGCCAGTGCATGTGACACAGG	3995
QY	3922	GACGAGAGGGCTGCGCCGCTGTGCTCCGCGCCAGTTCCTCTGCGCGGGGGTCACTGT	3981
Db	3996	GACGAGAGGGCTGCGCCGCTGTGCTCCGCGCCAGTTCCTCTGCGCGGGGGTCACTGT	4055
QY	3982	GTGACCTGTGCGCTGCGCGGAGGACGACTGTGACGACCGCTCAGACGAGGCGG	4041
Db	4056	GTGACCTGTGCGCTGCGCGGAGGACGACTGTGACGACCGCTCAGACGAGGCGG	4114
QY	4042	GACTGTGACGCGCATCTGCTCTGCCAAACAGTTCGGTGTGCGAGCGGCGAGTGTCTCTC	4101
Db	4116	GACTGTGACGCGCATCTGCTCTGCCAAACAGTTCGGTGTGCGAGCGGCGAGTGTCTCTC	4174
QY	4102	ATCAAAACGACGAGTGGACTCTTCTCCCGACTGTATCGACGCGCTCCGACGAGCTATGT	4161
Db	4176	ATCAAAACGACGAGTGGACTCTTCTCCCGACTGTATCGACGCGCTCCGACGAGCTATGT	4234
QY	4162	GAAATCAACGCGCCCTCAGACGACAGCCCGGCCCAACAGCAGTGCATCGGGCCCGTC	4221
Db	4236	GAAATCAACGCGCCCTCAGACGACAGCCCGGCCCAACAGCAGTGCATCGGGCCCGTC	4294
QY	4222	ATTGGCATCACT	4281
Db	4296	ATTGGCATCACT	4354
QY	4282	GTGTGCGACGCTATGCGGGGGCCAAACGCGCCCTTCCCGACGAGTATGTGACGCGGAC	4341
Db	4356	GTGTGCGACGCTATGCGGGGGCCAAACGCGCCCTTCCCGACGAGTATGTGACGCGGAC	4414
QY	4342	CCGACGCTGCGCCCTCAATTTTCATAGCCCGGGCGGTTCCACAGCATGGCCCTTACAGGC	4401
Db	4416	CCGACGCTGCGCCCTCAATTTTCATAGCCCGGGCGGTTCCACAGCATGGCCCTTACAGGC	4474
QY	4402	ATCGCATCGGAAAGTCCATGATGAGCTCGTGTAGCTGTATGGGGGGCGGGGGGGGGT	4461
Db	4476	ATCGCATCGGAAAGTCCATGATGAGCTCGTGTAGCTGTATGGGGGGCGGGGGGGGGT	4534
QY	4462	CCCCTGTACGACCGGAAACGAGTCAAGGGGCTCTGTCAGAGCTGTGTCAGACGAG	4521
Db	4536	CCCCTGTACGACCGGAAACGAGTCAAGGGGCTCTGTCAGAGCTGTGTCAGACGAG	4594
QY	4522	GCCACGCTGTACCGGCGGATCTCTGAAACCCCGCGCGCTCTCCCGGGCCACGGAACCCCTCCTG	4581

[illegible]

Search completed: February 18, 2005, 20:21:58
Job time : 811 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2005, 12:06:15 ; Search time 21176 Seconds
(without alignments)
11585.227 Million cell updates/sec

Title: US-09-931-375A-1
 Perfect score: 5063
 Sequence: 1 qccatgagcccccagatgac.....aggcctggagaaactttcta 5063

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

1:	gb_ba:
2:	gb_htg:
3:	gb_in:
4:	gb_om:
5:	gb_ov:
6:	gb_pat:
7:	gb_ph:
8:	gb_pl:
9:	gb_pr:
10:	gb_ro:
11:	gb_stss:
12:	gb_sy:
13:	gb_un:
14:	gb_vi:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES									
Result No.	Query	Score	Query			DB	ID	Description	
			Match	Length					
1	5061.4	100.0	5135	6	AX558229	AX558229	Sequence		
2	5061.4	100.0	5135	9	AF077820	AF077820	Homo sapi		
3	5058.2	99.9	5120	6	AX567475	AX567475	Sequence		
4	5058.2	99.9	5120	6	AX277522	AX277522	Sequence		
5	5058.2	99.9	5120	6	AX418086	AX418086	Sequence		
6	5056.8	99.9	5099	6	Q0870240	Q0870240	Sequence		
7	5056.6	99.9	5120	6	AX567476	AX567476	Sequence		
8	5056.6	99.9	5120	6	AX277523	AX277523	Sequence		
9	5056.6	99.9	5120	6	AX418087	AX418087	Sequence		
10	5052	99.8	5100	6	AX821900	AX821900	Sequence		
11	5052	99.8	5100	9	AF064548	AF064548	Homo sapi		
12	5028	99.3	5098	6	AX305075	AX305075	Sequence		
13	5028	99.3	5098	6	AX309179	AX309179	Sequence		
14	5028	99.3	5098	6	BD105986	BD105986	Novel LDL		
15	5010.8	99.0	5135	6	Q0870242	Q0870242	Sequence		
16	4999	98.7	5059	9	AB017498	AB017498	Homo sapi		
17	4902	96.8	5022	6	AX305086	AX305086	Sequence		
18	4902	96.8	5022	6	AX309190	AX309190	Sequence		
19	4902	96.8	5022	6	BD105997	BD105997	Novel LDL		

Db 218 CGGCTGTGGAGCGCGCGGAGTCAAGCTGGAGTCCACCATCGTGGTCAGCGGCTGGAG 277
Qy 241 GATGGCGCGAGTGGAGCTTCCAGTCTTTCAGGAGCGGTGTACTGACAGAGCTGAGC 300
Db 278 GATGGCGCGAGTGGAGCTTCCAGTCTTTCAGGAGCGGTGTACTGACAGAGCTGAGC 337
Qy 301 GAGGAGGCGCATCAAGCAGACCTACTGAAACAGACGGGGCGCGTGTGAGAACTGGTC 360
Db 338 GAGGAGGCGCATCAAGCAGACCTACTGAAACAGACGGGGCGCGTGTGAGAACTGGTC 397
Qy 361 ATCTCGGCGCTGGTCTTCCCGACGGGCTCGCTGCGACTGGTGGGCGAGAGCTGTAC 420
Db 398 ATCTCGGCGCTGGTCTTCCCGACGGGCTCGCTGCGACTGGTGGGCGAGAGCTGTAC 457
Qy 421 TGGACGGACTCAGAGCAACCGCATCGAGTGGCCAACTCAATGGCACATCCCGGAAG 480
Db 458 TGGACGGACTCAGAGCAACCGCATCGAGTGGCCAACTCAATGGCACATCCCGGAAG 517
Qy 481 GTGCTCTTCTGGCAGGACCTTGAACAGCCTTAGGGCCATCGGCTTGGACCCCGCTCACGGG 540
Db 518 GTGCTCTTCTGGCAGGACCTTGAACAGCCTTAGGGCCATCGGCTTGGACCCCGCTCACGGG 577
Qy 541 TACATGTACTGACAGACTGGGTGAGACGCCCGGATTGAGCGGCGAGGATGATGGC 600
Db 578 TACATGTACTGACAGACTGGGTGAGACGCCCGGATTGAGCGGCGAGGATGATGGC 637
Qy 601 AGCACCGGAAGATCAATGTGACTCGGACATTTACTGGCCCAATTGGACTGACCATCGAC 660
Db 638 AGCACCGGAAGATCAATGTGACTCGGACATTTACTGGCCCAATTGGACTGACCATCGAC 697
Qy 661 CTGGAGGACGAAAGCTTACTTGGGCTGACGCCAAGCTCAGCTTCAATCGAGTGGCCAA 720
Db 698 CTGGAGGACGAAAGCTTACTTGGGCTGACGCCAAGCTCAGCTTCAATCGAGTGGCCAA 757
Qy 721 CTGGAGGCTCTTCCGGCAGAGTGTGGAGGCGAGCTCAGCACACCCCTTCCGCTG 780
Db 758 CTGGAGGCTCTTCCGGCAGAGTGTGGAGGCGAGCTCAGCACACCCCTTCCGCTG 817
Qy 781 AGCGCTCTCCGGGACACTCTGTACTGGAAGAGTGGCAGACCGCTCCATCCATCCCTGC 840
Db 818 AGCGCTCTCCGGGACACTCTGTACTGGAAGAGTGGCAGACCGCTCCATCCATCCCTGC 877
Qy 841 AACAGCGCACTGGGGGGAAGAGAGAGATCTTGAGTGGCTTACTTCAACCCATGAC 900
Db 878 AACAGCGCACTGGGGGGAAGAGAGAGATCTTGAGTGGCTTACTTCAACCCATGAC 937
Qy 901 ATCCAGGTCTGAGCCAGGAGCGGAGCTTCTTCCACACTCGCTGTGAGGAGGACAA 960
Db 938 ATCCAGGTCTGAGCCAGGAGCGGAGCTTCTTCCACACTCGCTGTGAGGAGGACAA 997
Qy 961 GGGGCTGTCCCACTGTGCTGTGTCTTCCCAAGCGAGCTTCTTACACATGCGGCTGC 1020
Db 998 GGGGCTGTCCCACTGTGCTGTGTCTTCCCAAGCGAGCTTCTTACACATGCGGCTGC 1057
Qy 1021 CCCACGGGTGTGAGCTGAGGACAAAGAGAGAGTGAAGCAGGAGCGAGGAGGAGTG 1080
Db 1058 CCCACGGGTGTGAGCTGAGGACAAAGAGAGAGTGAAGCAGGAGCGGAGGAGGAGTG 1117
Qy 1081 CTGCTGTGCGCGCGGAGGAGCTTACGAGGAGTCTCGCTGAGACCGCGGACTTCAAC 1140
Db 1118 CTGCTGTGCGCGCGGAGGAGCTTACGAGGAGTCTCGCTGAGACCGCGGACTTCAAC 1177
Qy 1141 GACATCGTGTGAGGTGAGAGCATCTCGGACGCCATTTGCCATGACATACGACCGGCTA 1200
Db 1178 GACATCGTGTGAGGTGAGAGCATCTCGGACGCCATTTGCCATGACATACGACCGGCTA 1237
Qy 1201 GAGGGCTATGCTTACTGACAGATGAGGTGCGGCGCATCTCGAGGCGGTACTTGGAC 1260
Db 1238 GAGGGCTATGCTTACTGACAGATGAGGTGCGGCGCATCTCGAGGCGGTACTTGGAC 1297
Qy 1261 GGGTCTGGGCGGAGAGCTGTGTAACAGAGATCAAGCGACCCCGATGCGGCTC 1320

1298 GGGTCTGGGCGGAGAGCGTGTGTAACAGAGATCAAGCGACCCCGATGCGGCTC 1357
1321 GACTGGGTGGCGCGAAACCTCTACTGGAACGACACGGGCGACCGCATCGAGGTGACG 1380
1358 GACTGGGTGGCGCGAAACCTCTACTGGAACGACACGGGCGACCGCATCGAGGTGACG 1417
1381 CGCTCAAGCGGACCTTCCGCAAGATCTTGGTGTGCGAGGACCTGGAACGAGCCCGAGCC 1440
1418 CGCTCAAGCGGACCTTCCGCAAGATCTTGGTGTGCGAGGACCTGGAACGAGCCCGAGCC 1477
1441 ATCGCACTGACACCCCGTGTATGGGCTCATGTACTGGAACAGACTGGGAGAGAACCTTAAA 1500
1478 ATCGCACTGACACCCCGTGTATGGGCTCATGTACTGGAACAGACTGGGAGAGAACCTTAAA 1537
1501 ATCGAGTGTGCCAACTTGGATGGGAGGAGCGGCTGTGCTGTCAATGCTCCCTCGGG 1560
1538 ATCGAGTGTGCCAACTTGGATGGGAGGAGCGGCTGTGCTGTCAATGCTCCCTCGGG 1597
1561 TGGCCCAACCGCTTGGGCTTGGACCTGACGAGAGGGAAGCTCTACTGGGAGAGCGCAAG 1620
1598 TGGCCCAACCGCTTGGGCTTGGACCTGACGAGAGGGAAGCTCTACTGGGAGAGCGCAAG 1657
1621 ACAGACAAGATCGAGGTGATCAATGTGTGATGGAGAGGAGGAGCCCTCTCGGAGGAC 1680
1658 ACAGACAAGATCGAGGTGATCAATGTGTGATGGAGAGGAGGAGCCCTCTCGGAGGAC 1717
1681 AAGCTCCGCGACATTTTGGGTTTCCGCTGTGCGGAGCTTCACTACTTGGACTGACTGG 1740
1718 AAGCTCCGCGACATTTTGGGTTTCCGCTGTGCGGAGCTTCACTACTTGGACTGACTGG 1777
1741 CAGCGCGCAGCATCGAGCGGCTGCAAAAGCTCAAGCCAGCGGAGCGTCACTATTGAC 1800
1778 CAGCGCGCAGCATCGAGCGGCTGCAAAAGCTCAAGCCAGCGGAGCGTCACTATTGAC 1837
1801 CAGCTGCCCGACCTGATGGGCTCAAAAGCTGTGAATGTGGCAAGGCTGTGGAGCAAC 1860
1838 CAGCTGCCCGACCTGATGGGCTCAAAAGCTGTGAATGTGGCAAGGCTGTGGAGCAAC 1897
1861 CCGTGTGGGACAGAAACGGGGGTGAGCCAGCTGTGCTTCTTCAACCCCGAGCAAC 1920
1898 CCGTGTGGGACAGAAACGGGGGTGAGCCAGCTGTGCTTCTTCAACCCCGAGCAAC 1957
1921 CCGTGTGGCTCCCGCTTGGGCTTGGAGCTGTGAGTGAATGAAGAGCTGTGCT 1980
1958 CCGTGTGGCTCCCGCTTGGGCTTGGAGCTGTGAGTGAATGAAGAGCTGTGCT 2017
1981 GAGGCTTCTTGTGCTTTCACAGCAGAGCGGCTTCCAGAGATCTTCCCTCGAGACCAAT 2040
2018 GAGGCTTCTTGTGCTTTCACAGCAGAGCGGCTTCCAGAGATCTTCCCTCGAGACCAAT 2077
2041 AACACGAGCTGGGCTTCCCGCTCAAGGGCTTCAAGGAGGCTCAGCCCTCGACTTTGAT 2100
2078 AACACGAGCTGGGCTTCCCGCTCAAGGGCTTCAAGGAGGCTCAGCCCTGGACTTTGAT 2137
2101 GTGTCCAAACCAACCATCTACTGGAACAGCTGAGCTGGAAGACCATCAGCGGCGCTTC 2160
2138 GTGTCCAAACCAACCATCTACTGGAACAGCTGAGCTGGAAGACCATCAGCGGCGCTTC 2197
2161 ATGAACGGGAGCTCGGTGAGCAGCTGTGAGTGTGGCTTGTACTACCCCGAGGCGATG 2220
2198 ATGAACGGGAGCTCGGTGAGCAGCTGTGAGTGTGGCTTGTACTACCCCGAGGCGATG 2257
2221 GCGGTGTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAAGATTCGAA 2280
2258 GCGGTGTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAAGATTCGAA 2317
2281 GTGGCGGCTGGAGCGGCGAGTTCCGCAAGTCTCGCTGTGGAGGAGCTTGGACAAACCG 2340
2318 GTGGCGGCTGGAGCGGCGAGTTCCGCAAGTCTCGCTGTGGAGGAGCTTGGACAAACCG 2377
2341 AGGTCTGGCTTGGATCCCAAGGCTTACATCTACTTGGAGCGGCTTGGAGCGGCAAG 2400
2378 AGGTCTGGCTTGGATCCCAAGGCTTACATCTACTTGGAGCGGCTTGGAGCGGCAAG 2437

Db 578 TACATGTACTGACACAGACTGGGGTGAGACGCCCCGGATTGAGCGGCAGGGATGATGGC 637
Qy AGCACCCGGAGAGATCATTTGTCGACTCGGACATTTACTTGGCCCAATGGACTGACCATCGAC 660
Db 638 AGCACCCGGAGAGATCATTTGTCGACTCGGACATTTACTTGGCCCAATGGACTGACCATCGAC 697
Qy 661 CTGAGGAGCAGAAAGCTCTACTTGGGCTGACGCCAAGCTCAGCTTCATCCACCGTGCCCAAC 720
Db 698 CTGAGGAGCAGAAAGCTCTACTTGGGCTGACGCCAAGCTCAGCTTCATCCACCGTGCCCAAC 757
Qy 721 CTGACCGGCTCGTTCCGGCAGAAAGTGTGTGAGGCGAGCCCTGACGCAACCCCTTCCGCGTG 780
Db 758 CTGACCGGCTCGTTCCGGCAGAAAGTGTGTGAGGCGAGCCCTGACGCAACCCCTTCCGCGTG 817
Qy 781 ACGCTCTCCGGGACATCTCTACTGAGACAGACTGGCAGACCCGCTCCATCCATCCGCTGC 840
Db 818 ACGCTCTCCGGGACATCTCTACTGAGACAGACTGGCAGACCCGCTCCATCCATCCGCTGC 877
Qy 841 AACAAAGCGCACTGGGGGGAAGAGAGAGATCTCTGAGTGCCCTCTACTCAACCCATGGAC 900
Db 878 AACAAAGCGCACTGGGGGGAAGAGAGAGATCTCTGAGTGCCCTCTACTCAACCCATGGAC 937
Qy 901 ATCCAGGTGTGAGCCAGGAGCGGAGCGCTTTCTTCCACATCTGCTGTGAGGAGGACAAT 960
Db 938 ATCCAGGTGTGAGCCAGGAGCGGAGCGCTTTCTTCCACATCTGCTGTGAGGAGGACAAT 997
Qy 961 GCGGCTGTCTCCACCTGTGCTCTGCTGTCCCAAGCGAGCCCTTTTACACATGCGGCTGC 1020
Db 998 GCGGCTGTCTCCACCTGTGCTCTGCTGTCCCAAGCGAGCCCTTTTACACATGCGGCTGC 1057
Qy 1021 CCCACGGGTGTGAGCTGACGACAAACGGCAGGACGCTGAAGGAGAGCGGAGCGGAGGTG 1080
Db 1058 CCCACGGGTGTGAGCTGACGACAAACGGCAGGAGCTGTAAAGGAGGAGCGGAGGAGGTG 1117
Qy 1081 CTGCTGTGCGCCCGCGGAGCGGACCTTACGGAGGATCTGCTTGGACACGCGCGGACTTCAAC 1140
Db 1118 CTGCTGTGCGCCCGCGGAGCGGACCTTACGGAGGATCTGCTTGGACACGCGCGGACTTCAAC 1177
Qy 1141 GACATCGTGTGAGGTGGAGACGATCCGGCACGCCATTCGCACTGACACCGCGTA 1200
Db 1178 GACATCGTGTGAGGTGGAGACGATCCGGCACGCCATTCGCACTGACACCGCGTA 1237
Qy 1201 GAGGGCTATGTCTACTTGACAGATGACGAGGTGCGGGCCATCCGACGGGCGTACTCGAC 1260
Db 1238 GAGGGCTATGTCTACTTGACAGATGACGAGGTGCGGGCCATCCGACGGGCGTACTCGAC 1297
Qy 1261 GGGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGCGGTC 1320
Db 1298 GGGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGCGGTC 1357
Qy 1321 GACTGGGTGGCCGAAACCTCTACTTGGACGACACGCGGACGACCGGATCGAGGTGACG 1380
Db 1358 GACTGGGTGGCCGAAACCTCTACTTGGACGACACGCGGACGACCGGATCGAGGTGACG 1417
Qy 1381 CGCCTCAACGGCACCTCCGCAAGATCTGTGTGCGAGGACCTTGGAGCGGACCGGAGCC 1440
Db 1418 CGCCTCAACGGCACCTCCGCAAGATCTGTGTGCGAGGACCTTGGAGCGGACCGGAGCC 1477
Qy 1441 ATCCCACTGCAACCCGCTGATGGGCTCATGTACTTGGACAGACTGGGAGAGAACCTTAA 1500
Db 1478 ATCCCACTGCAACCCGCTGATGGGCTCATGTACTTGGACAGACTGGGAGAGAACCTTAA 1537
Qy 1501 ATCCAGGTGGCAACTTGGATGGGACGAGCGGCTGTGCTGGTCAATGCCCTCCCTCGG 1560
Db 1538 ATCCAGGTGGCAACTTGGATGGGACGAGCGGCTGTGCTGGTCAATGCCCTCCCTCGG 1597
Qy 1561 TGGCCCAACGGCTTGGCCCTGGACCTGCAGGAGGAGGAGCTTCTACTGGGAGAGCGCAAG 1620
Db 1598 TGGCCCAACGGCTTGGCCCTGGACCTGCAGGAGGAGGAGGAGCTTCTACTGGGAGAGCGCAAG 1657
Qy 1621 ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAAAGGCGGAGCCCTCTCTGGAGGAC 1680

Db 1658 ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCCCTCTCTGGAGAC 1717
Qy 1681 AAGCTCCCGCACATTTTCGGGTTCAAGCTGTCTGGGAGCTTCACTACTGACACTGACTGG 1740
Db 1718 AAGCTCCCGCACATTTTCGGGTTCAAGCTGTCTGGGAGCTTCACTACTGACACTGACTGG 1777
Qy 1741 CAGCGCCGACGATCGAGCGGGTGCAAAAGTCAAGGCGAGCCGCGGAGCTCATCATTTGAC 1800
Db 1778 CAGCGCCGACGATCGAGCGGGTGCAAAAGTCAAGGCGAGCCGCGGAGCTCATCATTTGAC 1837
Qy 1801 CAGCTGCCCGACCTGATGGGCTCAAAAGCTCAAGTGTGGAATGTGGCAAGGTCGTGGAAACCAAC 1860
Db 1838 CAGCTGCCCGACCTGATGGGCTCAAAAGCTCAAGTGTGGAATGTGGCAAGGTCGTGGAAACCAAC 1897
Qy 1861 CCGTGTGCGGACAGGACGGGGGTGAGCCACCTGTGCTTCTTCAACCCACCGCAAC 1920
Db 1898 CCGTGTGCGGACAGGACGGGGGTGAGCCACCTGTGCTTCTTCAACCCACCGCAAC 1957
Qy 1921 CCGTGTGCTGCCCCATCGGCTCGAGCTGTGAGTGACATGAAGACCTGCATCGTGCCT 1980
Db 1958 CCGTGTGCTGCCCCATCGGCTCGAGCTGTGAGTGACATGAAGACCTGCATCGTGCCT 2017
Qy 1981 GAGGCTTCTTGTGCTTCAACAGCAGAGCCGCCATCCACAGATCTCCCTCGAGACCAAT 2040
Db 2018 GAGGCTTCTTGTGCTTCAACAGCAGAGCCGCCATCCACAGATCTCCCTCGAGACCAAT 2077
Qy 2041 AACAAAGCTGGGCATCCGCTCAAGGGGTCAAGGAGGCTCAGCCCTGGACTTTGAT 2100
Db 2078 AACAAAGCTGGGCATCCGCTCAAGGGGTCAAGGAGGCTCAGCCCTGGACTTTGAT 2137
Qy 2101 GTGTCCAAACCAACATCTACTGACAGACGCTCAGCCCTGAAGACCATCAGCGCGGCTTC 2160
Db 2138 GTGTCCAAACCAACATCTACTGACAGACGCTCAGCCCTGAAGACCATCAGCGCGGCTTC 2197
Qy 2161 ATGAAACGGGAGCTCGGTGAGACGCTGTGAGTTTGGCTTGAATACCCGAGGGCATG 2220
Db 2198 ATGAAACGGGAGCTCGGTGAGACGCTGTGAGTTTGGCTTGAATACCCGAGGGCATG 2257
Qy 2221 GCGTGTGACTGGAATGGGCAAGAACCTCTACTGCGGCGGACACTGGGACCAACAGATCGAA 2280
Db 2258 GCGTGTGACTGGAATGGGCAAGAACCTCTACTGCGGCGGACACTGGGACCAACAGATCGAA 2317
Qy 2281 GTGCGCGGCTGGACGGGCGGATTCGGCAAGTCTCGTGTGAGGAGGACTTGGACAAACCG 2340
Db 2318 GTGCGCGGCTGGACGGGCGGATTCGGCAAGTCTCGTGTGAGGAGGACTTGGACAAACCG 2377
Qy 2341 AGGTGCTGCGGCTCGATGCCACCAAGGCTTACATCTAGACCGAGTGGGGGCGCAAG 2400
Db 2378 AGGTGCTGCGGCTCGATGCCACCAAGGCTTACATCTAGACCGAGTGGGGGCGCAAG 2437
Qy 2401 CCGAGGATCGTGGGCGCTTCATGGACGGGACCAACTGCACTGAGCGCTGGTGGACAGGTG 2460
Db 2438 CCGAGGATCGTGGGCGCTTCATGGACGGGACCAACTGCACTGAGCGCTGGTGGACAGGTG 2497
Qy 2461 GCGCGGCGCAACGACCTCACCATTTGACTACGCTGACCGAGCGCTTACTTGGACCGGCTG 2520
Db 2498 GCGCGGCGCAACGACCTCACCATTTGACTACGCTGACCGAGCGCTTACTTGGACCGGCTG 2557
Qy 2521 GACACCAACATGATCGAGTCTGTCGTCGTCGAGGAGCGGGTGTGATGGCGAC 2580
Db 2558 GACACCAACATGATCGAGTCTGTCGTCGTCGTCGAGGAGCGGGTGTGATGGCGAC 2617
Qy 2581 GATCTCCCGCACCGTTCCGTTCTGACGAGTACAGCGATTTATATCTACTTGGACAGACTGG 2640
Db 2618 GATCTCCCGCACCGTTCCGTTCTGACGAGTACAGCGATTTATATCTACTTGGACAGACTGG 2677
Qy 2641 AATCTGACAGCAATTGAGCGGCGGCGCAAGACTAGCGGCGGAAACCGCACCTCATTCAG 2700
Db 2678 AATCTGACAGCAATTGAGCGGCGGCGCAAGACTAGCGGCGGAAACCGCACCTCATTCAG 2737
Qy 2701 GGGCACCTGAGACTTGTGATGGACATCTCGTGTTCATCTCTCCCGCAGGATGGGCTC 2760
Db 2738 GGGCACCTGAGACTTGTGATGGACATCTCGTGTTCATCTCTCCCGCAGGATGGGCTC 2797

Qy	2761	AATGACTGTATGCAACAAAGGCGAGTGTGGGAGCTGTGCTTGGCATCCCGGGCGC	2820
Db	2798	AATGACTGTATGCAACAAAGGCGAGTGTGGGAGCTGTGCTTGGCATCCCGGGCGC	2857
Qy	2821	CACCGCTGCGGCTGCGCTCACACTACACCTTGACCCAGCAGCCGAACTGACAGCCG	2880
Db	2958	CACCGCTGCGGCTGCGCTCACACTACACCTTGACCCAGCAGCCGAACTGACAGCCG	2917
Qy	2881	CCCACCACTTCTGCTGTTTTCAGCCAGAAATCTGCGCATCTAGTCGGATGATCCCGAGAC	2940
Db	2918	CCCACCACTTCTGCTGTTTTCAGCCAGAAATCTGCGCATCTAGTCGGATGATCCCGAGAC	2977
Qy	2941	CAGCAGCCCGGATCTCATCTGCTCCCTGATGAGTGGAGGACGTCMAAGCCATCGAC	3000
Db	2978	CAGCAGCCCGGATCTCATCTGCTCCCTGATGAGTGGAGGACGTCMAAGCCATCGAC	3037
Qy	3001	TATGACCACTGGACAAAGTTCATCTACTGCGTGGATGGGCGCCAGAAATCAAGCGAGCC	3060
Db	3038	TATGACCACTGGACAAAGTTCATCTACTGCGTGGATGGGCGCCAGAAATCAAGCGAGCC	3097
Qy	3061	AAGGACGACGGGACCCAGCCCTTTGTTTGAACCTCTCTGAGCCCAAGGCCAATAACCGAGAC	3120
Db	3098	AAGGACGACGGGACCCAGCCCTTTGTTTGAACCTCTCTGAGCCCAAGGCCAATAACCGAGAC	3157
Qy	3121	AGGAGCCCGGACGACCTCAGCATCGACATCTACAGCCGGACACTGTTCTGACGCTGGAG	3180
Db	3158	AGGAGCCCGGACGACCTCAGCATCGACATCTACAGCCGGACACTGTTCTGACGCTGGAG	3217
Qy	3181	GCCACCAATACATCAACGCTCCAGAGCTGAGCGGGGAAAGCATGGGGGTGGTGTGCGT	3240
Db	3218	GCCACCAATACATCAACGCTCCAGAGCTGAGCGGGGAAAGCATGGGGGTGGTGTGCGT	3277
Qy	3241	GGGAGCCCGGACAGCCCGAGGCGCATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC	3300
Db	3278	GGGAGCCCGGACAGCCCGAGGCGCATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC	3337
Qy	3301	ACCAACATGACGAGCGGGGACGAAAGATCGAAGCGGAGCCCTGAGCGGACCGAGCGC	3360
Db	3338	ACCAACATGACGAGCGGGGACGAAAGATCGAAGCGGAGCCCTGAGCGGACCGAGCGC	3397
Qy	3361	GAGGTCTCTTCCACACCGGCTCATCTCGCCCTGTCGTCGTCGTCGTCGTCGTCGTCGTC	3420
Db	3398	GAGGTCTCTTCCACACCGGCTCATCTCGCCCTGTCGTCGTCGTCGTCGTCGTCGTCGTC	3457
Qy	3421	GGCAAGCTGTTCTGGGTGGAACGCGAACCTGGAAGCGCATTTGAGAGCTGTCAGGG	3480
Db	3458	GGCAAGCTGTTCTGGGTGGAACGCGAACCTGGAAGCGCATTTGAGAGCTGTCAGGG	3517
Qy	3481	GCCAAACGCTGACCTGAGGAGCGCAACATCTGTCAGGCTCTGGGCTGACCATCTT	3540
Db	3518	GCCAAACGCTGACCTGAGGAGCGCAACATCTGTCAGGCTCTGGGCTGACCATCTT	3577
Qy	3541	GGCAAGCATCTCTACTGATCGACCGGACGAGATGATCGAGCGTGTGGAGAGACC	3600
Db	3578	GGCAAGCATCTCTACTGATCGACCGGACGAGATGATCGAGCGTGTGGAGAGACC	3637
Qy	3601	ACCGGGGACAGCGGACCTGCGATCCAGGGCGTGTGCGGACCTCAGTCGGATCCATGCA	3660
Db	3638	ACCGGGGACAGCGGACCTGCGATCCAGGGCGTGTGCGGACCTCAGTCGGATCCATGCA	3697
Qy	3661	GTGGAGGAAGTCAGCTCGAGGAGTTCCTCAGCCCAACATGTCGCGGTGACAATGGTGGC	3720
Db	3698	GTGGAGGAAGTCAGCTCGAGGAGTTCCTCAGCCCAACATGTCGCGGTGACAATGGTGGC	3757
Qy	3721	TGCTCCCACTCTGATTTGCCAAGGGTGTGGGACCAACCGATGCTCATGCGGATCCAC	3780
Db	3758	TGCTCCCACTCTGATTTGCCAAGGGTGTGGGACCAACCGATGCTCATGCGGATCCAC	3817
Qy	3781	CTCGTGTCTTCGAGAACTGCTGACCTGTGGAGAGCGCCGACCTGCTCCCGGACCGAG	3840
Db	3818	CTCGTGTCTTCGAGAACTGCTGACCTGTGGAGAGCGCCGACCTGCTCCCGGACCGAG	3877

Qy	3841	TTTTCATGTGTCACAGGGAGATCGACTGTATCCCGGGGCGCTGGCGCTGTGACGGCTTT	3900
Db	3878	TTTTCATGTGTCACAGGGAGATCGACTGTATCCCGGGGCGCTGGCGCTGTGACGGCTTT	3937
Qy	3901	CCCGAGTGCATGATCCAGAGACGAGAGGGGTGCCCCGCTGCTCCGCGCCAGTTC	3960
Db	3938	CCCGAGTGCATGATCCAGAGACGAGAGGGGTGCCCCGCTGCTCCGCGCCAGTTC	3997
Qy	3961	CCCTGCGCGGGGTCACTGTGTGGACCTGGCTGCTGCGACGCGGAGGAGAGATGT	4020
Db	3998	CCCTGCGCGGGGTCACTGTGTGGACCTGGCTGCTGCGACGCGGAGGAGAGATGT	4057
Qy	4021	CAGGACCGCTCAGACGAGGGCGACTGTGACGCCATCTGCTGCGCCAAACAGTTCGGGT	4080
Db	4058	CAGGACCGCTCAGACGAGGGCGACTGTGACGCCATCTGCTGCGCCAAACAGTTCGGGT	4117
Qy	4081	GCGAGCGGCGAGTGTCTCTCATCAAAACAGAGTGGACCTCTTCCCGGACTGTATCGAC	4140
Db	4118	GCGAGCGGCGAGTGTCTCTCATCAAAACAGAGTGGACCTCTTCCCGGACTGTATCGAC	4177
Qy	4141	GGCTCCGACGAGCTCATGTGTGAAATCAACAGCGCCCTCAGACGACAGCCCGGCCAC	4200
Db	4178	GGCTCCGACGAGCTCATGTGTGAAATCAACAGCGCCCTCAGACGACAGCCCGGCCAC	4237
Qy	4201	AGCAGTGCCTCGGGCGGCTCATTTGGCATCATCTCTCTCTCTCTCTCTCTCTCTCTCT	4260
Db	4238	AGCAGTGCCTCGGGCGGCTCATTTGGCATCATCTCTCTCTCTCTCTCTCTCTCTCTCT	4297
Qy	4261	TATTTTGTGTGCGAGCGGTGTGCGAGCGCTATCGGGGGGCAACCGGGCCCTTCCCG	4320
Db	4298	TATTTTGTGTGCGAGCGGTGTGCGAGCGCTATCGGGGGGCAACCGGGCCCTTCCCG	4357
Qy	4321	CACGAGTATGTACGCGGAGCCCGCAGCTGCGCTCAATTTTCATAGCCCGGGGGTTC	4380
Db	4358	CACGAGTATGTACGCGGAGCCCGCAGCTGCGCTCAATTTTCATAGCCCGGGGGTTC	4417
Qy	4381	CAGCATGGCCCTTTTTCACAGGCTCGCATGCGGAAAGTCCATGATGAGTCCGTGAGCTG	4440
Db	4418	CAGCATGGCCCTTTTTCACAGGCTCGCATGCGGAAAGTCCATGATGAGTCCGTGAGCTG	4477
Qy	4441	ATGGGGGCGCGGGGCGGGGTGCTCTGTACGACCGGAAACACGTCACAGGGGCTTCTCC	4500
Db	4478	ATGGGGGCGCGGGGCGGGGTGCTCTGTACGACCGGAAACACGTCACAGGGGCTTCTCC	4537
Qy	4501	AGCAGCTCTCTCAGCAGCGGAGCCAGCTGTACCGCGGATCTCTGAAACCGCCCTTCC	4560
Db	4538	AGCAGCTCTCTCAGCAGCGGAGCCAGCTGTACCGCGGATCTCTGAAACCGCCCTTCC	4597
Qy	4561	CCGGCCACGAGCCCTCTCTGTACAAATGAGACATGTTTCTACTCTTCAAAATTCGGGCC	4620
Db	4598	CCGGCCACGAGCCCTCTCTGTACAAATGAGACATGTTTCTACTCTTCAAAATTCGGGCC	4657
Qy	4621	ACTCGGAGACCGTACAGCCCTCATCATTCGAGGAATGGCGCCCGGACGACGCGCTGC	4680
Db	4658	ACTCGGAGACCGTACAGCCCTCATCATTCGAGGAATGGCGCCCGGACGACGCGCTGC	4717
Qy	4681	AGCACCGAGCTGTGTGACGCGACTACAGCGCCAGCGCTGGAAGGCCAGCAAGTACTAC	4740
Db	4718	AGCACCGAGCTGTGTGACGCGACTACAGCGCCAGCGCTGGAAGGCCAGCAAGTACTAC	4777
Qy	4741	CTGATTTGAACTCGGACTCAGACCCCTATCCACCCCGACCAAGCCCGGACGAGTAC	4800
Db	4778	CTGATTTGAACTCGGACTCAGACCCCTATCCACCCCGACCAAGCCCGGACGAGTAC	4837
Qy	4801	CTGTGCGGGAGGACAGTGTGCCCTCTCGCCCGGACCGAGAGGAGTACTTCCATCTC	4860
Db	4838	CTGTGCGGGAGGACAGTGTGCCCTCTCGCCCGGACCGAGAGGAGTACTTCCATCTC	4897
Qy	4861	TTCCCGCCCTCTCGCTCCCTGCAAGGACTCATCTCTGACCTCGCGCCGGGCTCTGGC	4920
Db	4898	TTCCCGCCCTCTCGCTCCCTGCAAGGACTCATCTCTGACCTCGCGCCGGGCTCTGGC	4957
Qy	4921	TTCTCTGTGCCCCCTGTAAATAGTTTTTAAATATGAAACAAAGAAAAAATATATTTATGAT	4980

Db 1640 ACAGACAAAGATCGAGGTGATCAATGTTGATGGACGAAGAGCGGACCCCTCTCTGGAGGAC 1699
QY 1681 AAGCTCCCGCACATTTTCGGGTTACGCTGTCTGGGGGACTTCACTACTGAGACTGACTGG 1740
Db 1700 AAGCTCCCGCACATTTTCGGGTTACGCTGTCTGGGGGACTTCACTACTGAGACTGACTGG 1759
QY 1741 CAGCGCCGAGCATCGAGCGGGTGCACAAGGTCAAGGCGAGCGGGAGCTCATCATTCAC 1800
Db 1760 CAGCGCCGAGCATCGAGCGGGTGCACAAGGTCAAGGCGAGCGGGAGCTCATCATTCAC 1819
QY 1801 CAGCTGCCGACCTGATGGGCTCAAAAGCTGTGAATGTGGCCAAAGTCTCGGACCAAC 1860
Db 1820 CAGCTGCCGACCTGATGGGCTCAAAAGCTGTGAATGTGGCCAAAGTCTCGGACCAAC 1879
QY 1861 CCGTGTGGGACAGGAACGGGGGTGCAGCACCTGTGTCTTCTTACACCCCAACGCAACC 1920
Db 1880 CCGTGTGGGACAGGAACGGGGGTGCAGCACCTGTGTCTTCTTACACCCCAACGCAACC 1939
QY 1921 CCGTGTGGCTGCCCATCGGCTTGAGCTGTGTGAGTGACATGAAGACCTGCATCGTGCT 1980
Db 1940 CCGTGTGGCTGCCCATCGGCTTGAGCTGTGTGAGTGACATGAAGACCTGCATCGTGCT 1999
QY 1981 GAGGCTTCTTGGTCTTCAACAGACAGACCGCCATCCACAGGATCTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTTCTTGGTCTTCAACAGACAGACCGCCATCCACAGGATCTCCCTCGAGACCAAT 2059
QY 2041 AACACGAGCTGGCCATCCCGCTCAGCGGCGTCAAGGAGGCTCAGCCCTGGACTTTGAT 2100
Db 2060 AACACGAGCTGGCCATCCCGCTCAGCGGCGTCAAGGAGGCTCAGCCCTGGACTTTGAT 2119
QY 2101 GTGTCCAAACCAACCATCTACTGACAGACGTGAGCCTGAAGACCATCAGCGCGGCTTC 2160
Db 2120 GTGTCCAAACCAACCATCTACTGACAGACGTGAGCCTGAAGACCATCAGCGCGGCTTC 2179
QY 2161 ATGAACGGGAGTCTGGTGGAGACGCTGTGAGTTTGGCTTGACTACCCCGAGGGCATG 2220
Db 2180 ATGAACGGGAGTCTGGTGGAGACGCTGTGAGTTTGGCTTGACTACCCCGAGGGCATG 2239
QY 2221 GCCGTGAGCTGGAGGGAAGAACTTACTTCTGTCGGCGGACACTGGGACCAACAGATCGAA 2280
Db 2240 GCCGTGAGCTGGAGGGAAGAACTTACTTCTGTCGGCGGACACTGGGACCAACAGATCGAA 2299
QY 2281 GTGGCGGCTGGAGCGGCACTTCGGCAAGTCTCTGTGTGGAGGGACTTGGAACAACCG 2340
Db 2300 GTGGCGGCTGGAGCGGCACTTCGGCAAGTCTCTGTGTGGAGGGACTTGGAACAACCG 2359
QY 2341 AGGTGCTGGGCTGGATCCCAACAAAGGGCTACATCTACTGGAACGAGTGGGGGGCAAG 2400
Db 2360 AGGTGCTGGGCTGGATCCCAACAAAGGGCTACATCTACTGGAACGAGTGGGGGGCAAG 2419
QY 2401 CCGAGGATCGTGGGCTTTCATGGACGGGACCAACTGCATGACGCTGGTGGACAAAGTG 2460
Db 2420 CCGAGGATCGTGGGCTTTCATGGACGGGACCAACTGCATGACGCTGGTGGACAAAGTG 2479
QY 2461 GCGCGGGCAACGACCTCACCATTGACTACGCTGACAGCGCTTACTTGGACCGACCTG 2520
Db 2480 GCGCGGGCAACGACCTCACCATTGACTACGCTGACAGCGCTTACTTGGACCGACCTG 2539
QY 2521 GACACCAACATGATGAGTGTGTCAAACATGTGGGTGAGGAGCGGGTCTGATTCGCGAC 2580
Db 2540 GACACCAACATGATGAGTGTGTCAAACATGTGGGTGAGGAGCGGGTCTGATTCGCGAC 2599
QY 2581 GATCTCCGCAACCGCTTCGCTGACGAGTACAGCGATTATNCTACTGACAGACTGG 2640
Db 2600 GATCTCCGCAACCGCTTTCGCTGACGAGTACAGCGATTATNCTACTGACAGACTGG 2659
QY 2641 AATCTGCAACGATTTGAGCGGCGGACAAAGACTAGCGGCGGGAACCGCACCTCATCCAG 2700
Db 2660 AATCTGCAACGATTTGAGCGGCGGACAAAGACTAGCGGCGGGAACCGCACCTCATCCAG 2719
QY 2701 GGCACCTTGGATTCGTGATGACATCTCTGTGTGTTCATCTCTCCGCGAGGATGGCCTC 2760
Db 2720 GGCACCTTGGATTCGTGATGACATCTCTGTGTGTTCATCTCTCCGCGAGGATGGCCTC 2779

QY 2761 AATGACTGTATGCACAAACGGGCACTGTGGGCACTGTGCCTTGCCATCCCCGGCGC 2820
Db 2780 AATGACTGTATGCACAAACGGGCACTGTGGGCACTGTGCCTTGCCATCCCCGGCGC 2839
QY 2821 CACCGCTGCGGCTGCGCTCACACTACACCTTGACACCCAGCAGCGCAACTGACAGCCG 2880
Db 2840 CACCGCTGCGGCTGCGGCTCACACTACACCTTGACACCCAGCAGCGCAACTGACAGCCG 2899
QY 2881 CCCACCACTTCTTGTGCTTTCAGCCAGAAATCTGCTCATCTGATCGGATGATCCCGAGCAC 2940
Db 2900 CCCACCACTTCTTGTGCTTTCAGCCAGAAATCTGCTCATCTGATCGGATGATCCCGAGCAC 2959
QY 2941 CAGCACAGCCCGGATCTCATCTCCCTGCTGATGAGTGAAGACGTCAAGACCATCCAC 3000
Db 2960 CAGCACAGCCCGGATCTCATCTCCCTGCTGATGAGTGAAGACGTCAAGACCATCCAC 3019
QY 3001 TATGACCCACCTGGACAAAGTTTCATCTACTGGGTGATGGGCGCCAGAAACATCAAGCGAGCC 3060
Db 3020 TATGACCCACCTGGACAAAGTTTCATCTACTGGGTGATGGGCGCCAGAAACATCAAGCGAGCC 3079
QY 3061 AAGGACAGCGGGAACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAAAGGCCAAACCCAGAC 3120
Db 3080 AAGGACAGCGGGAACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAAAGGCCAAACCCAGAC 3139
QY 3121 AGGACAGCCCAACGACCTCAGCATCGACATCTACAGCGGACACTGTTCTGGACGTGCGAG 3180
Db 3140 AGGACAGCCCAACGACCTCAGCATCGACATCTACAGCGGACACTGTTCTGGACGTGCGAG 3199
QY 3181 GCCACCAATACATCAACAGCTCCACAGGCTGACGCGGGAAGCCATGGGGGTGTGTCGCT 3240
Db 3200 GCCACCAATACATCAACAGCTCCACAGGCTGACGCGGGAAGCCATGGGGGTGTGTCGCT 3259
QY 3241 GGGGACCGCGACAAGCCAGGCGCATCTGCTGTAACGGCGAGCGAGGCTACTGTFACCTTC 3300
Db 3260 GGGGACCGCGACAAGCCAGGCGCATCTGCTGCTCAACGGCGAGCGAGGCTACTGTFACCTTC 3319
QY 3301 ACCAACATGACAGGACCGGGACGCAAGATCGAACCGCAGCCCTGGAGCGCACCGAGCGC 3360
Db 3320 ACCAACATGACAGGACCGGGACGCAAGATCGAACCGCAGCCCTGGAGCGCACCGAGCGC 3379
QY 3361 GAGGTCTCTTTCACCAACCGGCTCATCCGCCCTGTGGCCCTGTGGTGGAGCAACACACTG 3420
Db 3380 GAGGTCTCTTTCACCAACCGGCTCATCCGCCCTGTGGCCCTGTGGTGGAGCAACACACTG 3439
QY 3421 GGCAGAGCTTCTGGGTGGAGCGGACCTGAGCGCATTTGAGAGCTGTGACTGTCAAGG 3480
Db 3440 GGCAGAGCTTCTGGGTGGAGCGGACCTGAGCGCATTTGAGAGCTGTGACTGTGTCAGGG 3499
QY 3481 GCCAACCGCTGACCTTGAGGAGCGCAACATCTGTGACGCTCTGGGCTTGACCATCTCT 3540
Db 3500 GCCAACCGCTGACCTTGAGGAGCGCAACATCTGTGACGCTCTGGGCTTGACCATCTCT 3559
QY 3541 GGCAGAGCTCTTACTGATCGACCGCGCAGCAGAGATGATCGAGCGTGTGGAGAGACC 3600
Db 3560 GGCAGAGCTCTTACTGATCGACCGCGCAGCAGAGATGATCGAGCGTGTGGAGAGACC 3619
QY 3601 ACCGGGAGCAAGCGGACTCGGCATCCAGGGCGGTGCGCCACCTCACTGGCATCCATGCA 3660
Db 3620 ACCGGGAGCAAGCGGACTCGGCATCCAGGGCGGTGCGCCACCTCACTGGCATCCATGCA 3679
QY 3661 GTGGAGGAAGTCAAGCTGGAGGAGTCTTACGCCACCCATGTGCCCGTGAACAATGGTGGC 3720
Db 3680 GTGGAGGAAGTCAAGCTGGAGGAGTCTTACGCCACCCATGTGCCCGTGAACAATGGTGGC 3739
QY 3721 TGCTCCCAACATCTGATTTGCAAGGGTGAATGGGACACACCGTGTCTATGCCCAGTCCAC 3780
Db 3740 TGCTCCCAACATCTGATTTGCAAGGGTGAATGGGACACACCGTGTCTATGCCCAGTCCAC 3799
QY 3781 CTGCTGCTCTGCAAGAACCTGCTGACCTGTGGAGGCGGCCACCTGCTCCCGGACGAG 3840
Db 3800 CTGCTGCTCTGCAAGAACCTGCTGACCTGTGGAGGCGGCCACCTGCTCCCGGACGAG 3859

Db 500 GTGCTCTTCTGGCAGGACCTTTGACAGCGAGGGCCATCGCCTTGGACCCCGCTCAACGGG 559
QY 541 TACATGTAATGACACAGACTGGGGTGAAGCGCCCGGATTTAGCGGGGAGGATGATGGC 600
Db 560 TACATGTAATGACACAGACTGGGGTGAAGCGCCCGGATTTAGCGGGGAGGATGATGGC 619
QY 601 AGCACCCGAGATCATTTGTGACATCGGACATTTACTGGCCCAATGGACTGACCATCGAC 660
Db 620 AGCACCCGAGATCATTTGTGACATCGGACATTTACTGGCCCAATGGACTGACCATCGAC 679
QY 661 CTGGAGGAGCAAGACTCTTACTGGGCTGACGCAAGCTCAGCTTCAATCCACCGTGCCCAAC 720
Db 680 CTGGAGGAGCAAGACTCTTACTGGGCTGACGCAAGCTCAGCTTCAATCCACCGTGCCCAAC 739
QY 721 CTGGAGGAGCAAGACTCTTACTGGGCTGACGCAAGCTCAGCTTCAATCCACCGTGCCCAAC 780
Db 740 CTGGAGGAGCAAGACTCTTACTGGGCTGACGCAAGCTCAGCTTCAATCCACCGTGCCCAAC 799
QY 781 ACCTCTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCCGCTCCATCCATGSCCTGC 840
Db 800 ACCTCTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCCGCTCCATCCATGSCCTGC 859
QY 841 AACAGCCCACTGGGGGAGAGGAGAGATCTCTGAGTGCCTCTACTCAACCCATGGAC 900
Db 860 AACAGCCCACTGGGGGAGAGGAGAGATCTCTGAGTGCCTCTACTCAACCCATGGAC 919
QY 901 ATCCAGGTGCTGAGCCAGGAGCGGAGCTTTCTTCCACACTCGCTGTGAGAGGACAAAT 960
Db 920 ATCCAGGTGCTGAGCCAGGAGCGGAGCTTTCTTCCACACTCGCTGTGAGAGGAGGACAAAT 979
QY 961 GCGGCTCTCCCACTCTGCTGCTGCCCAAGCAGGCTTTCTACACATGCGCCTGC 1020
Db 980 GCGGCTCTCCCACTCTGCTGCTGCCCAAGCAGGCTTTCTACACATGCGCCTGC 1039
QY 1021 CCCAGGTGTGAGCTGACAGCAAAACGCGAGAGCTGTAAAGGAGGAGCGGAGAGGTG 1080
Db 1040 CCCAGGTGTGAGCTGACAGCAAAACGCGAGAGCTGTAAAGGAGGAGCGGAGAGGTG 1099
QY 1081 CTGCTGTGCGCCGCGGAGCGGAGCTTACGAGGAGTCTCGCTGGACACGCGGAGCTTAC 1140
Db 1100 CTGCTGTGCGCCGCGGAGCGGAGCTTACGAGGAGTCTCGCTGGACACGCGGAGCTTAC 1159
QY 1141 GACATCGTGTGCAAGGTGAGACATCCGCGACGCAATGCGCATCGACTACGACCCGCTA 1200
Db 1160 GACATCGTGTGCAAGGTGAGACATCCGCGACGCAATGCGCATCGACTACGACCCGCTA 1219
QY 1201 GAGGCTATGTCTACTGACAGATGACGAGTGGGGCCATCCGAGGGCGTACTGGAC 1260
Db 1220 GAGGCTATGTCTACTGACAGATGACGAGTGGGGCCATCCGAGGGCGTACTGGAC 1279
QY 1261 GGGTCTGGGGCGAGAGCTGGTCAACACGAGATCAACGACCCCGATGGCATCGCGTC 1320
Db 1280 GGGTCTGGGGCGAGAGCTGGTCAACACGAGATCAACGACCCCGATGGCATCGCGTC 1339
QY 1321 GACTGGGTGGCCGAAACCTCTACTGGACGACACGCGGACCGGATCAGAGTGACG 1380
Db 1340 GACTGGGTGGCCGAAACCTCTACTGGACGACACGCGGACCGGATCAGAGTGACG 1399
QY 1381 GCGCTCAACGCGACCTCCGCAAGATCTGTGTGCGAGGACCTGGAGAGCCCGGAGCC 1440
Db 1400 GCGCTCAACGCGACCTCCGCAAGATCTGTGTGCGAGGACCTGGAGAGCCCGGAGCC 1459
QY 1441 ATCGCACTGACCCCGCTGATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAA 1500
Db 1460 ATCGCACTGACCCCGCTGATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAA 1519
QY 1501 ATCGAGTGTGCAACTTGGATGGGAGGAGCGGCTGTGTGCTCAATGCTCCCTCGG 1560
Db 1520 ATCGAGTGTGCAACTTGGATGGGAGGAGCGGCTGTGTGCTCAATGCTCCCTCGG 1579
QY 1561 TGGCCCAACGCGCTGGCTGACCTCGAGAGGAGGAGCTTACTTGGGAGAGCGCAAG 1620
Db 1580 TGGCCCAACGCGCTGGCTGACCTCGAGAGGAGGAGCTTACTTGGGAGAGCGCAAG 1639

QY 1621 ACAGCAAGATCGAGGTGATCAATGTTGATGGAGCAAGAGCGGACCCCTCTCTGAGGAC 1680
Db 1640 ACAGCAAGATCGAGGTGATCAATGTTGATGGAGCAAGAGCGGACCCCTCTCTGAGGAC 1699
QY 1681 AAGCTCCCGCACATTTTCGGGTTCACGCTGCTGGGGACTTCATCTACTGAGCTGACTGG 1740
Db 1700 AAGCTCCCGCACATTTTCGGGTTCACGCTGCTGGGGACTTCATCTACTGAGCTGACTGG 1759
QY 1741 CAGCGCCGACGATCGAGCGGGTGCAAGAGTCAAGGCGAGCCGCGGACGTCATCATTTGAC 1800
Db 1760 CAGCGCCGACGATCGAGCGGGTGCAAGAGTCAAGGCGAGCCGCGGACGTCATCATTTGAC 1819
QY 1801 CAGCTGCCCGACCTGATGGGGCTCAAGCTGTGAAATGTGGCAAGAGTGTGGAACCAAC 1860
Db 1820 CAGCTGCCCGACCTGATGGGGCTCAAGAGTGTGAAATGTGGCAAGAGTGTGGAACCAAC 1879
QY 1861 CCGTGTGCGGACAGGAAACGGGGGTGACAGCCACTGTGCTTTTTCACACCCACGCAAC 1920
Db 1880 CCGTGTGCGGACAGGAAACGGGGGTGACAGCCACTGTGCTTTTTCACACCCACGCAAC 1939
QY 1921 CCGTGTGCTGCCCCCATCGGCTGAGCTGTGAGTGAATGAAGACCTGTCATCGTGCCT 1980
Db 1940 CCGTGTGCTGCCCCCATCGGCTGAGCTGTGAGTGAATGAAGACCTGTCATCGTGCCT 1999
QY 1981 GAGGCTTCTTGTGCTTTCACGAGCAGAGCGGCTCCACAGGATCTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTTCTTGTGCTTTCACGAGCAGAGCGGCTCCACAGGATCTCCCTCGAGACCAAT 2059
QY 2041 AACAAACGAGTGGGCTATCCGCTCAAGGGCTCAAGAGGCTCAGCCCTGAGCTTTGAT 2100
Db 2060 AACAAACGAGTGGGCTATCCGCTCAAGGGCTCAAGAGGCTCAGCCCTGAGCTTTGAT 2119
QY 2101 GTGTCCAAACCAACATCTACTGACAGAGCTGAGCTGAAAGACATGAGCGGCGCTTC 2160
Db 2120 GTGTCCAAACCAACATCTACTGACAGAGCTGAGCTGAAAGACATGAGCGGCGCTTC 2179
QY 2161 ATGAAACGAGGCTCGGTGGAGCAGTGTGGAGTTTGGCTTGTACTACCCCGAGGGCATG 2220
Db 2180 ATGAAACGAGGCTCGGTGGAGCAGTGTGGAGTTTGGCTTGTACTACCCCGAGGGCATG 2239
QY 2221 GCCCTTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2280
Db 2240 GCCCTTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2299
QY 2281 GTGCGCGGCTGAGCGGCGAGTTCCGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCG 2340
Db 2300 GTGCGCGGCTGAGCGGCGAGTTCCGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCG 2359
QY 2341 AGGTGCTGGCCCTGGATCCCAAGGCTTACATCTACTGGACCGGCTGTGAGCGGCAAG 2400
Db 2360 AGGTGCTGGCCCTGGATCCCAAGGCTTACATCTACTGGACCGGCTGTGAGCGGCAAG 2419
QY 2401 CCGAGGATCGTGGGGCTTTCATGGACGGGACCAACTGCAATGCTGTGGAGCAAGGTG 2460
Db 2420 CCGAGGATCGTGGGGCTTTCATGGACGGGACCAACTGCAATGCTGTGGAGCAAGGTG 2479
QY 2461 GCGCGGCGCAACGACCTTCAATTTGACTACGTCAGCAGCGCTTCTACTGAGCAGCCTG 2520
Db 2480 GCGCGGCGCAACGACCTTCAATTTGACTACGTCAGCAGCGCTTCTACTGAGCAGCCTG 2539
QY 2521 GACACCAACATGATCGAGTGTGTCCAAATGCTGGGTGAGGCGGGTGTGATTTGCCGAC 2580
Db 2540 GACACCAACATGATCGAGTGTGTCCAAATGCTGGGTGAGGCGGGTGTGATTTGCCGAC 2599
QY 2581 GATCTCCCGCACCGGCTGTGAGCGGATGACGAGTATATATCTACTGAGCAGCTGG 2640
Db 2600 GATCTCCCGCACCGGCTGTGAGCGGATGACGAGTATATATCTACTGAGCAGCTGG 2659
QY 2641 AATCTGACAGCATTGAGCGGCGGCAAGACTAGCGGCGGACCCGACCTCATCCAG 2700
Db 2660 AATCTGACAGCATTGAGCGGCGGCGCAAGACTAGCGGCGGCGGAAACCCGACCTCATCCAG 2719

Qy	2701	GGCCACCTGGACTTTCGTGATGGACATCCTCGTGTGTTCCACTCTCTCCGCGAGGATGGCGTC	2761
Db	2720	GGCCACCTGGACTTTCGTGATGGACATCCTCGTGTGTTCCACTCTCTCCGCGAGGATGGCGTC	2779
Qy	2761	AATGACTGTATGCACAAACACGGGACAGTGTGGGACAGCTGTGSCCTTGGCCATCTCCCGCGGCGC	2820
Db	2780	AATGACTGTATGCACAAACACGGGACAGTGTGGGACAGCTGTGSCCTTGGCCATCTCCCGCGGCGC	2839
Qy	2821	CACCGCTCGCGCTCGCGCTCACTACACCTTGGACCCCGACAGCCGCAACTGCGAGTATCCCGGACGAC	2880
Db	2840	CACCGCTCGCGCTCGCGCTCACTACACCTTGGACCCCGACAGCCGCAACTGCGAGTATCCCGGACGAC	2899
Qy	2881	CCCAACCACTTCTTGTCTGTTCAGCCAGAGAAATCTGCGATCAGTCCGATGATCCCGGACGAC	2940
Db	2900	CCCAACCACTTCTTGTCTGTTCAGCCAGAGAAATCTGCGATCAGTCCGATGATCCCGGACGAC	2959
Qy	2941	CAGCACAGCCGAGTCTCATCTCTGCCCTGTGATGGAATGAGGAAAGCTCAAGAGCCATCGAC	3000
Db	2960	CAGCACAGCCGAGTCTCATCTCTGCCCTGTGATGGAATGAGGAAAGCTCAAGAGCCATCGAC	3019
Qy	3001	TATGACCCACTGGACAAGTTTCATCTACTCTGGTGGATGGCGCCAGAGAACATCAAGCGAGCC	3060
Db	3020	TATGACCCACTGGACAAGTTTCATCTACTCTGGTGGATGGCGCCAGAGAACATCAAGCGAGCC	3079
Qy	3061	AAGSACAGCGGACCCAGCCCTTGTGTTTGTGACTCTCTGAGCCAGAGCCGCAAAACCCAGAC	3120
Db	3080	AAGSACAGCGGACCCAGCCCTTGTGTTTGTGACTCTCTGAGCCAGAGCCGCAAAACCCAGAC	3139
Qy	3121	AGGCAGCCCAAGACCTCAGATCGACATCTACAGCCGAGACATGTTCTCTGAGCGTGGAG	3180
Db	3140	AGGCAGCCCAAGACCTCAGATCGACATCTACAGCCGAGACATGTTCTCTGAGCGTGGAG	3199
Qy	3181	GCCACCAATACCATCAACGCTCCACAGGCTGAGCGGGGAAGCCATGGGGGTGGTCTGCTCGT	3240
Db	3200	GCCACCAATACCATCAACGCTCCACAGGCTGAGCGGGGAAGCCATGGGGGTGGTCTGCTCGT	3259
Qy	3241	GGGGACCGGACAAGCCACAGGCCCACATCGTCTGTCAAACGCGGAGCGAGGGTACTGTACTTC	3300
Db	3260	GGGGACCGGACAAGCCACAGGCCCACATCGTCTGTCAAACGCGGAGCGAGGGTACTGTACTTC	3319
Qy	3301	ACCACATGACAGACCGGGCAGCCCAAGATCGAACGCGAGCCCTGAGAGCGACCGAGCGC	3360
Db	3320	ACCACATGACAGACCGGGCAGCCCAAGATCGAACGCGAGCCCTGAGAGCGACCGAGCGC	3379
Qy	3361	GAGTCTCTTTCACCAACCGGCTCATCCGCGCTGTGGCCCTGTGTGGTGGACAACAACACTG	3420
Db	3380	GAGTCTCTTTCACCAACCGGCTCATCCGCGCTGTGGCCCTGTGTGGTGGACAACAACACTG	3439
Qy	3421	GGCAAGCTGTTTGGGTGGACGGGACCTGAAGCGCAATGAGAGCTGTGACTGTCAAGG	3480
Db	3440	GGCAAGCTGTTTGGGTGGACGGGACCTGAAGCGCAATGAGAGCTGTGACTGTCAAGG	3499
Qy	3481	GCCACCGCTGACCTGGAGGAGCCCAACATCGTGGAGCGCTCTGGGCTGTGACATCTT	3540
Db	3500	GCCACCGCTGACCTGGAGGAGCCCAACATCGTGGAGCGCTCTGGGCTGTGACATCTT	3559
Qy	3541	GGCAAGCATCTCTACTGGAATCGACCGGACGACGATGATCGAGCGTGTGGAGAGAGCC	3600
Db	3560	GGCAAGCATCTCTACTGGAATCGACCGGACGACGATGATCGAGCGTGTGGAGAGAGCC	3619
Qy	3601	ACCGGGACAAGCGGACTCGCATCCAGGGCCGTGTGCCCACTCACTGGCATCCCATGCA	3660
Db	3620	ACCGGGACAAGCGGACTCGCATCCAGGGCCGTGTGCCCACTCACTGGCATCCCATGCA	3679
Qy	3661	GTGGAGGAAGTCAAGCTGGAGGAGTTCTCAGCCCAACCATGTGCCCTGTGACATGTGGC	3720
Db	3680	GTGGAGGAAGTCAAGCTGGAGGAGTTCTCAGCCCAACCATGTGCCCTGTGACATGTGGC	3739
Qy	3721	TGCTCCCAACATCTGATTTCGCAAGGGGTGATGGGACACCAAGGTGCTCATSCCAGATCCAC	3780
Db	3740	TGCTCCCAACATCTGATTTCGCAAGGGGTGATGGGACACCAAGGTGCTCATSCCAGATCCAC	3799
Qy	3781	CTCGTGTCTCTGCAAGACCTGTGACTGTGGAGAGCGCGCCACCTGCTCTCCCGGACCGAG	3840

[illegible]

Qy	361	ATCTCCGGCCTGGTCTCTCCCGACGGCCTCGCCTGCGACTGGGTGGGCAAGAAGCTGTAC	420
Db	380	ATCTCCGGCCTGGTCTCTCCCGACGGCCTCGCCTGCGACTGGGTGGGCAAGAAGCTGTAC	439
Qy	421	TGGACGGACTCAGAGACAAACCGCATCGAGTGGCCAACTCAATGGCACATCCCGGAG	480
Db	440	TGGACGGACTCAGAGACAAACCGCATCGAGTGGCCAACTCAATGGCACATCCCGGAG	499
Qy	481	GTGCTCTTCTGGCAGGACCTTGGACCGAGCTTAGGGCCATCGCCTTTGGACCCCGCTCACGGG	540
Db	500	GTGCTCTTCTGGCAGGACCTTGGACCGAGCTTAGGGCCATCGCCTTTGGACCCCGCTCACGGG	559
Qy	541	TACATGTACTGGACAGACTGGGTGGAGCGCCCGGATTGAGCGGGCAGGATGGATGGC	600
Db	560	TACATGTACTGGACAGACTGGGTGGAGCGCCCGGATTGAGCGGGCAGGATGGATGGC	619
Qy	601	AGCACCCGGGAAGATCATTTGTGGACTCGGACATTTACTTGGCCCAATGGACTGACCATCGAC	660
Db	620	AGCACCCGGGAAGATCATTTGTGGACTCGGACATTTACTTGGCCCAATGGACTGACCATCGAC	679
Qy	661	CTGAGAGGACAGAAGCTCTACTGGGCTGACGCCAAAGCTCAGCTTCAATCCACGTCGCAAC	720
Db	680	CTGAGAGGACAGAAGCTCTACTGGGCTGACGCCAAAGCTCAGCTTCAATCCACGTCGCAAC	739
Qy	721	CTGACCGGCTCGTTCCGGCAAGTGGTGGAGGCGACCTTGCACGCCCTTCGCCCTG	780
Db	740	CTGACCGGCTCGTTCCGGCAAGTGGTGGAGGCGACCTTGCACGCCCTTCGCCCTG	799
Qy	781	ACGCTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCAATCCATGCTGTC	840
Db	800	ACGCTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCAATCCATGCTGTC	859
Qy	841	AACAAAGCGACTGGGGGGAAGAGAGAGACTCTGAGTGCCCTCTACTCACCAATGGAC	900
Db	860	AACAAAGCGACTGGGGGGAAGAGAGAGACTCTGAGTGCCCTCTACTCACCAATGGAC	919
Qy	901	ATCCAGGTGCTGAGCCAGGAGCGCAGCTTTTCTTCCACACTCGCTGTGAGGAGGACAAT	960
Db	920	ATCCAGGTGCTGAGCCAGGAGCGCAGCTTTTCTTCCACACTCGCTGTGAGGAGGACAAT	979
Qy	961	GGCGGCTGCTCCACCTGTGCTGCTGTCCCAAGCAGAGCCTTTCTACACATGGCGCTGC	1020
Db	980	GGCGGCTGCTCCACCTGTGCTGCTGTCCCAAGCAGAGCCTTTCTACACATGGCGCTGC	1039
Qy	1021	CCACGGGTGTGACGTGTCAGGACAAACGGCAGGACGTGTAAAGCAGAGCCGAGGAGGTG	1080
Db	1040	CCACGGGTGTGACGTGTCAGGACAAACGGCAGGACGTGTAAAGCAGAGCCGAGGAGGTG	1099
Qy	1081	CTGCTGTGCGCCGGCGGACGGACCTTAGGAGATCTCGCTGGACACGCCGGAATTCAAC	1140
Db	1100	CTGCTGTGCGCCGGCGGACGGACCTTAGGAGATCTCGCTGGACACGCCGGAATTCAAC	1159
Qy	1141	GACATCTGTCTGAGGTGGACGACATCCGGCAGCCCAATGGCCATCGACTGACGCCGCTA	1200
Db	1160	GACATCTGTCTGAGGTGGACGACATCCGGCAGCCCAATGGCCATCGACTGACGCCGCTA	1219
Qy	1201	GAGGGCTATGTCTACTGGAACAGATGACGAGGTGGCGCCATCCGACGGGCGTACTGAGC	1260
Db	1220	GAGGGCTATGTCTACTGGAACAGATGACGAGGTGGCGCCATCCGACGGGCGTACTGAGC	1279
Qy	1261	GGGTCTGGGCGCAGACGCTGTCTCAACACCGAGATCAACGACCCCGATGGCATCGCGTTC	1320
Db	1280	GGGTCTGGGCGCAGACGCTGTCTCAACACCGAGATCAACGACCCCGATGGCATCGCGTTC	1339
Qy	1321	GACTGGGTGGCCGAAACCTTACTTGGACCGACACGGGCGACGGACCGCATCGAGGTGACG	1380
Db	1340	GACTGGGTGGCCGAAACCTTACTTGGACCGACACGGGCGACGGACCGCATCGAGGTGACG	1399
Qy	1381	CGCCTCAACGGCACCTCCGCAAGATCCTGTGTGCGAGACCTTGGACGAGCCCCGAGCC	1440
Db	1400	CGCCTCAACGGCACCTCCGCAAGATCCTGTGTGCGAGACCTTGGACGAGCCCCGAGCC	1459
Qy	1441	ATCGCACTGCACCCCGTGTATGGGCGCTCATGTACTTGGACAGACTGGGAGAGAACCCCTAAA	1500

2540	Db	GACACCAACATGATCGAGTCTGCTCAACATGCTGGGTGAGGAGCGGGTCTGTGATTGGCCGAC	2599
2581	Qy	GATCTCCGCGACCCGTTGGGTCTGACGCGAGTACGCGATTATATCTACTTGGACAGACTGG	2640
2600	Db	GATCTCCGCGACCCGTTTCGGTCTGACGCGAGTACGCGATTATATCTACTTGGACAGACTGG	2659
2641	Qy	AATCTGCACAGCANTTGACGGGGCCGACAAGACTAGCGCGCGGAACCGGACCCCTCATCCAG	2700
2660	Db	AATCTGCACAGCANTTGACGGGGCCGACAAGACTAGCGCGCGGAACCGGACCCCTCATCCAG	2719
2701	Qy	GGCCACCTTGGACTTCTGTATGCGACATCTCTGTGTTCACATCTCTCCGCGAGGATGGCCCT	2760
2720	Db	GGCCACCTTGGACTTCTGTATGCGACATCTCTGTGTTCACATCTCTCCGCGAGGATGGCCCT	2779
2761	Qy	AATGACTGTATGCAACAACCGGCGAGTGTGGCGAGCTGTGCTTGGCATCCCCGCGCGGC	2820
2780	Db	AATGACTGTATGCAACAACCGGCGAGTGTGGCGAGCTGTGCTTGGCATCCCCGCGCGGC	2839
2821	Qy	CACGCTGCGGCTGGGCTCTACATACACCTTGACCCCGAGCGCCGAGCGGAACTGCGAGCCCG	2880
2840	Db	CACGCTGCGGCTGGGCTCTACATACACCTTGACCCCGAGCGCCGAGCGGAACTGCGAGCCCG	2899
2881	Qy	CCCAACCACTTCTGCTGTTTCAGCCGAGAAATCTGCCATCAGTCCGATGATCCCGGAGCGAC	2940
2900	Db	CCCAACCACTTCTGCTGTTTCAGCCGAGAAATCTGCCATCAGTCCGATGATCCCGGAGCGAC	2959
2941	Qy	CAGCACAGCCCGGATCTCATCTCTGCCCTGCAATGGAATGAGGAACGTCAAAGCCATCGAC	3000
2960	Db	CAGCACAGCCCGGATCTCATCTCTGCCCTGCAATGGAATGAGGAACGTCAAAGCCATCGAC	3019
3001	Qy	TATGACCACTGGAACAAGTTCAATCTACTGGTGGATGGGCGCCAGAAATCAACGAGCGACC	3060
3020	Db	TATGACCACTGGAACAAGTTCAATCTACTGGTGGATGGGCGCCAGAAATCAACGAGCGACC	3079
3061	Qy	AAGGACAGCGGACCGACCCCTTTGTTTTCACCTCTCTGAGCCGAGGCCAAACCCAGAC	3120
3080	Db	AAGGACAGCGGACCGACCCCTTTGTTTTCACCTCTCTGAGCCGAGGCCAAACCCAGAC	3139
3121	Qy	AGGCAGCGCCACGACCTCAGCATCGACATCTACAGCCGCGACACTGTTCTGGAAGTGGCGAG	3180
3140	Db	AGGCAGCGCCACGACCTCAGCATCGACATCTACAGCCGCGACACTGTTCTGGAAGTGGCGAG	3199
3181	Qy	GCCACCAATACCATCAACGTTCCAGAGCTGAGCGGGGAAGCCATGGGGGTGGTCTGCGT	3240
3200	Db	GCCACCAATACCATCAACGTTCCAGAGCTGAGCGGGGAAGCCATGGGGGTGGTCTGCGT	3259
3241	Qy	GGGACCGCGACAGCCCGAGGCCATCTGTGTCTAAACGCGAGCGAGGGTACTGTACTTC	3300
3260	Db	GGGACCGCGACAGCCCGAGGCCATCTGTGTCTAAACGCGAGCGAGGGTACTGTACTTC	3319
3301	Qy	ACCAACATGAGGACCGGGCAGCCAAAGATCGAACGCGCAGCCCTGGAGCGCAGCCGAGCGC	3360
3320	Db	ACCAACATGAGGACCGGGCAGCCAAAGATCGAACGCGCAGCCCTGGAGCGCAGCCGAGCGC	3379
3361	Qy	GAGTCTCTTTCACACCGGCTCATCTCGGCCCTTGGCCCTTGGTGTGGGACCAACACTG	3420
3380	Db	GAGTCTCTTTCACACCGGCTCATCTCGGCCCTTGGCCCTTGGTGTGGGACCAACACTG	3439
3421	Qy	GGCAAGCTGTTCTGGGTGGAACGCGAATCGAAGCGCAATGAGAGCTGTGACCTGTCAAGG	3480
3440	Db	GGCAAGCTGTTCTGGGTGGAACGCGAATCGAAGCGCAATGAGAGCTGTGACCTGTCAAGG	3499
3481	Qy	GCCAAACGCTGACCTTGGAGGACGCGCAATCTGTGAGCCTCTGGGCCCTGACCATCTT	3540
3500	Db	GCCAAACGCTGACCTTGGAGGACGCGCAATCTGTGAGCCTCTGGGCCCTGACCATCTT	3559
3541	Qy	GGCAAGCATCTCTACTTGGATCGACCGCAGCAGCAGATGATCGAGCGTGTGGAGAGACCC	3600
3560	Db	GGCAAGCATCTCTACTTGGATCGACCGCAGCAGCAGATGATCGAGCGTGTGGAGAGACCC	3619
3601	Qy	ACGGGGAACAAGCGGACTCGCATCCAGGGCGGCTGTCGCCCACTCACTCGCATCCATGCA	3660
3620	Db	ACGGGGAACAAGCGGACTCGCATCCAGGGCGGCTGTCGCCCACTCACTCGCATCCATGCA	3679

QY	3661	GTGAGGAGAGTACGCTTGGAGGAGTTCTTCAGGCCACCCATGTGCGCGTGACAATGTGTGGC	3721
DB	3680	GTGAGGAAGTACGCTTGGAGGAGTTCTTCAGGCCACCCATGTGCGCGTGACAATGTGTGGC	3739
QY	3721	TGCTCCACACATCTGTATTGGCCAGGGGTGATGGGACACACGCTGCTCATGCCAGTCCAC	3780
DB	3740	TGCTCCACACATCTGTATTGGCCAGGGGTGATGGGACACACGCTGCTCATGCCAGTCCAC	3799
QY	3781	CTCGTGTCTCTGTCAGAAACCTGTGACCTGTGGAGAGCGGCCACCTGCTCTCCCGGACAG	3840
DB	3800	CTCGTGTCTCTGTCAGAAACCTGTGACCTGTGGAGAGCGGCCACCTGCTCTCCCGGACAG	3859
QY	3841	TTTGGATGTGCGCACAGGGGAGATCGACTGTATATCCCGGGGCTGTGGCGCTGTGACGGCTTT	3900
DB	3860	TTTGGATGTGCGCACAGGGGAGATCGACTGTATATCCCGGGGCTGTGGCGCTGTGACGGCTTT	3919
QY	3901	CCCGAGTGCATGACACAGAGCGACGAGGGGCTGCCCGTGTGCTTCCGCGGCCAGTTC	3960
DB	3920	CCCGAGTGCATGACACAGAGCGACGAGGGGCTGCCCGTGTGCTTCCGCGGCCAGTTC	3979
QY	3961	CCCTGCGCGGGGTCAGTGTGTGGACCTTCGGCTTCGCTGCGACGCGCAGGGCAGACTGT	4020
DB	3980	CCCTGCGCGGGGTCAGTGTGTGGACCTTCGCTGCGCTGCGACGCGCAGGGCAGACTGT	4039
QY	4021	CAGGACCGCTCAGACGAGGCGGACTGTGACGCCATCTGCTGCGCCACCACTGCTCCGGTGT	4080
DB	4040	CAGGACCGCTCAGACGAGGCGGACTGTGACGCCATCTGCTGCGCCACCACTGCTCCGGTGT	4099
QY	4081	GCAGCGGCGAGTGTCTCATCAACACAGAGTGTGGACTCTTCCCGACTGTATTCGAC	4140
DB	4100	GCAGCGGCGAGTGTCTCATCAACACAGAGTGTGGACTCTTCCCGACTGTATTCGAC	4159
QY	4141	GGCTCCGACGAGCTCATGTGTGAAATCACCAAGCGCCCTCAGACGACAGCCCGGCCAC	4200
DB	4160	GGCTCCGACGAGCTCATGTGTGAAATCACCAAGCGCCCTCAGACGACAGCCCGGCCAC	4219
QY	4201	AGCAGTGCCATCGGGCCCGCTCATTTGGCATCATCTCTCTCTCTTCTGTCATGGGTGTGTC	4260
DB	4220	AGCAGTGCCATCGGGCCCGCTCATTTGGCATCATCTCTCTCTCTTCTGTCATGGGTGTGTC	4279
QY	4261	TATTTTGTGTGCACGCGGTGTGTGCGAGCGTATGTGGGGGCGCAACGGGCGCTTCCCG	4320
DB	4280	TATTTTGTGTGCACGCGGTGTGTGCGAGCGTATGTGGGGGCGCAACGGGCGCTTCCCG	4339
QY	4321	CACGAGTATGTCAAGCGGACCCCGCAGTGCCTCTCAATTTTCATAGCCCGGCGGTTC	4380
DB	4340	CACGAGTATGTCAAGCGGACCCCGCAGTGCCTCTCAATTTTCATAGCCCGGCGGTTC	4399
QY	4381	CAGCATGGCCCTTTCACAGGCATCGCATGCGGAAAGTCCATGATGAGTCCGTGAGCGTG	4440
DB	4400	CAGCATGGCCCTTTCACAGGCATCGCATGCGGAAAGTCCATGATGAGTCCGTGAGCGTG	4459
QY	4441	ATGGGGGCGGGGCGGGGTGCCCTGTATCGAACCGGAACCACTGTCAGGGGCGCTCGTCC	4500
DB	4460	ATGGGGGCGGGGCGGGGTGCCCTGTATCGAACCGGAACCACTGTCAGGGGCGCTCGTCC	4519
QY	4501	AGCAGCTGTGTACAGCAGAAAGGCGAGCTGTACCGCCGATCTCTGAAACCGCGCGCCCTCC	4560
DB	4520	AGCAGCTGTGTACAGCAGAAAGGCGAGCTGTACCGCCGATCTCTGAAACCGCGCGCCCTCC	4579
QY	4561	CCGGCCACGAGCCCTCCCTGTGTACACATGGACATGTTCTTCTTCTTCAACATTCGGGCC	4620
DB	4580	CCGGCCACGAGCCCTCCCTGTGTGTACACATGGACATGTTCTTCTTCTTCAACATTCGGGCC	4639
QY	4621	ACTGCGAGCCGTACAGGCGCTTACATTTTCGAGGAATGCGCGCCCGCCGACGAGCGCCCTGC	4680
DB	4640	ACTGCGAGCCGTACAGGCGCTTACATTTTCGAGGAATGCGCGCCCGCCGACGAGCGCCCTGC	4699
QY	4681	AGCACCGAGTGTGTGACAGGACTACAGCGCCAGCGCGTGTGGAAGGCGCAAGTACTAC	4740
DB	4700	AGCACCGAGTGTGTGACAGGACTACAGCGCGCGTGTGGAAGGCGCAAGTACTAC	4759

Db 1400 GCGCTCAACGCGCACCTCCGCAAGATCTGGTGTGCGAGGACCTGGACGAGCCCGGAGCC 1459
QY 1441 ATCGACATGCAACCCCGTGTATGGGCCCTCATGTACTGGACAGACTGGGGAGAGAAACCTTAAA 1500
Db 1460 ATCGACATGCAACCCCGTGTATGGGCCCTCATGTACTGGACAGACTGGGGAGAGAAACCTTAAA 1519
QY 1501 ATCGAGTGTGCCCACTCTGGATGGGAGGAGCGGCGTGTGGTCAATGCCCTCCCTCGGG 1560
Db 1520 ATCGAGTGTGCCCACTCTGGATGGGAGGAGCGGCGTGTGGTCAATGCCCTCCCTCGGG 1579
QY 1561 TGGCCCAACGCGCTGGCCCTGGACCTGACAGAGGGGAAGCTCTACTGGGGAGAGCCCAAG 1620
Db 1580 TGGCCCAACGCGCTGGCCCTGGACCTGACAGAGGGGAAGCTCTACTGGGGAGAGCCCAAG 1639
QY 1621 ACAGACAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCCCTCTGGAGGAC 1680
Db 1640 ACAGACAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCCCTCTGGAGGAC 1699
QY 1681 AAGCTCCCGCAATTTTCGGGTTACGCTGTCTGGGGGACTTTCATCTACTGGACTGACTGG 1740
Db 1700 AAGCTCCCGCAATTTTCGGGTTACGCTGTCTGGGGGACTTTCATCTACTGGACTGACTGG 1759
QY 1741 CAGCGCCGACGATCGAGCGGGTGCACAAAGGTCAAGGCCAGCGGGAAGTCTCATATTGAC 1800
Db 1760 CAGCGCCGACGATCGAGCGGGTGCACAAAGGTCAAGGCCAGCGGGAAGTCTCATATTGAC 1819
QY 1801 CAGCTGCCCGACCTGATGGGGCTCAAGCTGTGTAATGTGGCCAAAGGTCTGTCGGAACCAAC 1860
Db 1820 CAGCTGCCCGACCTGATGGGGCTCAAGCTGTGTAATGTGGCCAAAGGTCTGTCGGAACCAAC 1879
QY 1861 CCGTGTGCGGACAGAAACGGGGGTGACGACCTGTGCTTTCACACCCCGACGCAACC 1920
Db 1880 CCGTGTGCGGACAGAAACGGGGGTGACGACCTGTGCTTTCACACCCCGACGCAACC 1939
QY 1921 CCGTGTGCGCTGCCCATCGGCTCGAGCTGTGAGTGCATGAAGACCTGTCATCGTGCT 1980
Db 1940 CCGTGTGCGCTGCCCATCGGCTCGAGCTGTGAGTGCATGAAGACCTGTCATCGTGCT 1999
QY 1981 GAGGCTTTCTTGGTCTTACAGAGAGCGCGCATCCACAGATCTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTTTCTTGGTCTTACAGAGAGCGCGCATCCACAGATCTCCCTCGAGACCAAT 2059
QY 2041 AACACGACGTGGCCATCCCGTCAACGCGGCTCAAGGAGGCTCAGCCCTGCACTTTGAT 2100
Db 2060 AACACGACGTGGCCATCCCGTCAACGCGGCTCAAGGAGGCTCAGCCCTGCACTTTGAT 2119
QY 2101 GTGTCCAAACACCATCTACTGGAACAGCTGAGCCCTGAAGACCATCAGCCGGCGCTTC 2160
Db 2120 GTGTCCAAACACCATCTACTGGAACAGCTGAGCCCTGAAGACCATCAGCCGGCGCTTC 2179
QY 2161 ATGAACGGGAGCTCGGTGGAGCACTGTGTGAGTTTGGCCCTTGACTACCCGAGGGCATG 2220
Db 2180 ATGAACGGGAGCTCGGTGGAGCACTGTGTGAGTTTGGCCCTTGACTACCCGAGGGCATG 2239
QY 2221 GCCGTGACTGATGGGCAAGAACTTACTTGGGCGGACACTGGGACCAACAGATCGAA 2280
Db 2240 GCCGTGACTGATGGGCAAGAACTTACTTGGGCGGACACTGGGACCAACAGATCGAA 2299
QY 2281 GTGGCGGCGCTGGACGGGCGAGTTCCGGCAAGTCTCTGTGTGGAGGACTTGGACAAACCCG 2340
Db 2300 GTGGCGGCGCTGGACGGGCGAGTTCCGGCAAGTCTCTGTGTGGAGGACTTGGACAAACCCG 2359
QY 2341 AGGTGCTGGCCCTGGATCCCAACGAGGCTTACTTACTGGAACGAGTGGGGCGGCAAG 2400
Db 2360 AGGTGCTGGCCCTGGATCCCAACGAGGCTTACTTACTGGAACGAGTGGGGCGGCAAG 2419
QY 2401 CCGAGGATCTGCGGGCTTTCATGGAACGGGACCACTGATGACGCTGGTGGACAAAGTG 2460
Db 2420 CCGAGGATCTGCGGGCTTTCATGGAACGGGACCACTGATGACGCTGGTGGACAAAGTG 2479
QY 2461 GGCCGGGCAACGACCTCACCATTGACTACGCTGACAGCGCTTCTACTGGACGACCTG 2520

Db 2480 GGCCGGGCAACGACCTTCACATTGACTACGCTGACAGCGCTCTACTGGACCGACCTG 2539
QY 2521 GACACCAACATGATCGAGTCTGTCACCAATGCTGGGTGAGGAGCGGTCTGTATTGCCGAC 2580
Db 2540 GACACCAACATGATCGAGTCTGTCACCAATGCTGGGTGAGGAGCGGTCTGTATTGCCGAC 2599
QY 2581 GATCTCCCGCACCCGTTCTGACGCGAGTACAGCGATTATATCTACTGGACAGACTGG 2640
Db 2600 GATCTCCCGCACCCGTTCTGACGCGAGTACAGCGATTATATCTACTGGACAGACTGG 2659
QY 2641 AATCTGCAACAGCATGTAGCGGCGGACAAAGACTAGCGGCGGAAACCGACCTCATCCAG 2700
Db 2660 AATCTGCAACAGCATGTAGCGGCGGACAAAGACTAGCGGCGGAAACCGACCTCATCCAG 2719
QY 2701 GGCCACCTGGACTTTGTTGATGGACATCTCTGCTGTTTCCATCTCTCCCGCAGGATGGCTC 2760
Db 2720 GGCCACCTGGACTTTGTTGATGGACATCTCTGCTGTTTCCATCTCTCCCGCAGGATGGCTC 2779
QY 2761 AATGACTGTATGCAACAAACGCGGAGTGTGGGAGCTGTGCTTGGCCATCCCCGGCGGC 2820
Db 2780 AATGACTGTATGCAACAAACGCGGAGTGTGGGAGCTGTGCTTGGCCATCCCCGGCGGC 2839
QY 2821 CACCGCTGCGGCTCGGCTCTACACTACACCTTGGACCCAGCAGCCGAACTGCAAGCCG 2880
Db 2840 CACCGCTGCGGCTCGGCTCTACACTACACCTTGGACCCAGCAGCCGAACTGCAAGCCG 2899
QY 2881 CCAACGACCTTCTTGTGTTTTCAGCCAGAAATCTGCGCATCAGTCGAGTATCCCGGACGAC 2940
Db 2900 CCAACGACCTTCTTGTGTTTTCAGCCAGAAATCTGCGCATCAGTCGAGTATCCCGGACGAC 2959
QY 2941 CAGCAGAGCCCGGATCTCATCTCCCTGATGAGTGTGAGGAAAGCTCAAAGCCATCGAC 3000
Db 2960 CAGCAGAGCCCGGATCTCATCTCCCTGATGAGTGTGAGGAAAGCTCAAAGCCATCGAC 3019
QY 3001 TATGACCCCACTGGACAAAGTTTCATCTAGTGGTGGATGGCGGCCAGAAATCAAGCGAGCC 3060
Db 3020 TATGACCCCACTGGACAAAGTTTCATCTAGTGGTGGATGGCGGCCAGAAATCAAGCGAGCC 3079
QY 3061 AAGCAGAGCGGACCCAGCCCTTGTGTTTGTGACCTCTCTGAGCCAAAGGCCAAACCCAGAC 3120
Db 3080 AAGCAGAGCGGACCCAGCCCTTGTGTTTGTGACCTCTCTGAGCCAAAGGCCAAACCCAGAC 3139
QY 3121 AGGCGAGCCCAACGACCTCAGCATCGACATCTACAGCGGACACTGTTCTTGACGCTGCGAG 3180
Db 3140 AGGCGAGCCCAACGACCTCAGCATCGACATCTACAGCGGACACTGTTCTTGACGCTGCGAG 3199
QY 3181 GCCACCAATACCATCAACGCTCCAGAGCTGACGCGGGGAAGCCATGGGGGTGTTGCTGCT 3240
Db 3200 GCCACCAATACCATCAACGCTCCAGAGCTGACGCGGGGAAGCCATGGGGGTGTTGCTGCT 3259
QY 3241 GGGGACCGGACCAAGCCAGGGCCATCGTGTGTAACGCGGAGCGAGGGTACTGTACTTTC 3300
Db 3260 GGGGACCGGACCAAGCCAGGGCCATCGTGTGTAACGCGGAGCGAGGGTACTGTACTTTC 3319
QY 3301 ACCAAACATGACGAGCCGCGGACCAAGATCGAAACGCGAGCCCTGGAGCGGACCCAGCGC 3360
Db 3320 ACCAAACATGACGAGCCGCGGACCAAGATCGAAACGCGAGCCCTGGAGCGGACCCAGCGC 3379
QY 3361 GAGTCTCTTTCACCAACCGGCTCATCCGCCCTGTGCGCTTGGTGGTGGACCAACACTG 3420
Db 3380 GAGTCTCTTTCACCAACCGGCTCATCCGCCCTGTGCGCTTGGTGGTGGACCAACACTG 3439
QY 3421 GGAAGCTGTTCTGGGTGGACCGGACCTGGAAGCGCATTTGAGAGCTGTGACTGTCAAGG 3480
Db 3440 GGAAGCTGTTCTGGGTGGACCGGACCTGGAAGCGCATTTGAGAGCTGTGACTGTCAAGG 3499
QY 3481 GCCAAACCGCTGACCTTGGAGGACGCGCAACATCTGTGAGCTCTGGGCTTCACCATCTT 3540
Db 3500 GCCAAACCGCTGACCTTGGAGGACGCGCAACATCTGTGAGCTCTGGGCTTCAGCATCTT 3559
QY 3541 GGCAAGCATCTTACTGATCGACCGGACGAGCAGATGATCGAGCGCTGTGGAGAGACC 3600
Db 3560 GGCAAGCATCTTACTGATCGACCGGACGAGCAGATGATCGAGCGCTGTGGAGAGACC 3619

[illegible]

Db	1321	TGGGTGCGCCGAAACCTCTTACTGGACCGACA	CGGGGACCGACCGCATCTGAGGTGACGCGC	1381
Qy	1384	CTCAACGGCACCTCCCGCAAGATCTCGGTGT	TCGGAGGACCTTGGACGAGCCCCGAGCCATC	1443
Db	1381	CTCAACGGCACCTCCCGCAAGATCTCGGTGT	TCGGAGGACCTTGGACGAGCCCCGAGCCATC	1440
Qy	1444	GCACTGCACCCCGTGATGGCGCTCATGTA	CTGCACAGACTCTGGGAGAGAACCTTAAATC	1503
Db	1441	GCACTGCACCCCGTGATGGCGCTCATGTA	CTGCACAGACTCTGGGAGAGAACCTTAAATC	1500
Qy	1504	GAGTGTGCCAACTTGGATGGGACGAGCGCGT	GTCTGTCTCAATGCTCCCTCGGGTGG	1563
Db	1501	GAGTGTGCCAACTTGGATGGGACGAGCGCGT	GTCTGTCTCAATGCTCCCTCGGGTGG	1560
Qy	1564	CCCAACGGCCCTGGCCCTTGGACCTTGCAG	AGGGGGAAGCTTACTTGGGAGACGCAAGACA	1623
Db	1561	CCCAACGGCCCTGGCCCTTGGACCTTGCAG	AGGGGGAAGCTTACTTGGGAGACGCAAGACA	1620
Qy	1624	GACAAGATTCGAGGTGATCAATGTTGAT	TGGGACGAAAGGCGGACCTTCTTGGAGGACAAG	1681
Db	1621	GACAAGATTCGAGGTGATCAATGTTGAT	TGGGACGAAAGGCGGACCTTCTTGGAGGACAAG	1680
Qy	1684	CTCCCGCACATTTTCGGGTTTCA	CGCTGCTGGGGGACTTCTACTTGGACTGACTGCGAG	1743
Db	1681	CTCCCGCACATTTTCGGGTTTCA	CGCTGCTGGGGGACTTCTACTTGGACTGACTGCGAG	1740
Qy	1744	CGCGCAGCATTCGAGCGGGTGCA	CAAGGCTCAAGGCGGACCGGACGTCATCTATTGACCAG	1803
Db	1741	CGCGCAGCATTCGAGCGGGTGCA	CAAGGCTCAAGGCGGACCGGACGTCATCTATTGACCAG	1800
Qy	1804	CTGCGCAGCTGATGGGGCTCAAAGCTGT	GAAATGTGTGGCCAAAGTCTGTGCGAAACCAACCCG	1863
Db	1801	CTGCGCAGCTGATGGGGCTCAAAGCTGT	GAAATGTGTGGCCAAAGTCTGTGCGAAACCAACCCG	1860
Qy	1864	TGTGCGACAGGAACGGGGGGTTCAGCGCA	CCCTGTGCTTCTTCAACCCACGCGCAACCCGG	1923
Db	1861	TGTGCGACAGGAACGGGGGGTTCAGCGCA	CCCTGTGCTTCTTCAACCCACGCGCAACCCGG	1920
Qy	1924	TGTGGCTGCCCCCATCTGGCCCTGGAGCT	GTCTCAGTGACAATGAAGACCTGTGATCTGTCCTTGAG	1983
Db	1921	TGTGGCTGCCCCCATCTGGCCCTGGAGCT	GTCTCAGTGACAATGAAGACCTGTGATCTGTCCTTGAG	1980
Qy	1984	GCCTTCTTGCTTTCACAGCAGAGCGCCAT	CCACAGGATCTCCCTTCGAGACCAATPAAC	2043
Db	1981	GCCTTCTTGCTTTCACAGCAGAGCGCCAT	CCACAGGATCTCCCTTCGAGACCAATPAAC	2040
Qy	2044	AACGACGTGGCCATCCCGCTCA	CGGGGCTCAAGGAGGCTTCAGCCCTTGGACTTGTG	2103
Db	2041	AACGACGTGGCCATCCCGCTCA	CGGGGCTCAAGGAGGCTTCAGCCCTTGGACTTGTG	2100
Qy	2104	TCCAAACACACATCTACTGACAGAGCT	CAGCTGAGACCATCAAGCCGCGCTTCCATG	2163
Db	2101	TCCAAACACACATCTACTGACAGAGCT	CAGCTGAGACCATCAAGCCGCGCTTCCATG	2160
Qy	2164	AACGGGAGCTCGGTGGAGCAGCTGGTGG	AGCTTTTGGCTTGAATCAACCCGAGGCGATGGCC	2223
Db	2161	AACGGGAGCTCGGTGGAGCAGCTGGTGG	AGCTTTTGGCTTGAATCAACCCGAGGCGATGGCC	2220
Qy	2224	GTTTGACTGGATGGGCAAGAACCTTCTACT	TGGGCGGACACTGGGACCAACAGATCGAAGTG	2283
Db	2221	GTTTGACTGGATGGGCAAGAACCTTCTACT	TGGGCGGACCACTGGGACCAACAGATCGAAGTG	2280
Qy	2284	GCGCGGCTGCAGCGGCAAGTTTCCGCGA	AGTCTCGTGTGGAGGAGACTTGGACAACCCGAGG	2343
Db	2281	GCGCGGCTGCAGCGGCAAGTTTCCGCGA	AGTCTCGTGTGGAGGAGACTTGGACAACCCGAGG	2340
Qy	2344	TCGCTGGCCCTTGGATCCCAACAAAGGG	CTACATCTACTTGGACCGAGTGGGGCGGCAAGCCG	2403
Db	2341	TCGCTGGCCCTTGGATCCCAACAAAGGG	CTACATCTACTTGGACCGAGTGGGGCGGCAAGCCG	2400
Qy	2404	AGGATCTGCGGGCTTTCATGGA	CGGGACCAACTGCATGACGCTGTGTGGA	2463
Db	2401	AGGATCTGCGGGCTTTCATGGA	CGGGACCAACTGCATGACGCTGTGTGGA	2460

QY 2464 CGGGCAACGACCTCAACGATTGATACGCTGACAGCGCCTCTACTGAGCGACCTGGAC 2523
DB |||||
2461 CGGGCAACGACCTCAACGATTGATACGCTGACAGCGCCTCTACTGAGCGACCTGGAC 2520
QY |||||
DB |||||
2524 ACCAATCATGATCGAGTCGTCGAACATGCTGGGTCTAGGAGCGGTCGTGATTCGCGAGAT 2583
DB |||||
2521 ACCAATCATGATCGAGTCGTCGAACATGCTGGGTCTAGGAGCGGTCGTGATTCGCGAGAT 2580
QY |||||
DB |||||
2584 CTCGCGCACCCGTTCCGTTCTGACGCGAGTACAGCGATTATATCTACTGAGCAGACTGGAT 2643
DB |||||
2581 CTCGCGCACCCGTTCCGTTCTGACGCGAGTACAGCGATTATATCTACTGAGCAGACTGGAT 2640
QY |||||
DB |||||
2644 CTGCAAGCATTTAGCGGGCCGACAGACTAGCGCGGGAACCGCACCTCTATTCAGAGGC 2703
DB |||||
2641 CTGCAAGCATTTAGCGGGCCGACAGACTAGCGCGGGAACCGCACCTCTATTCAGAGGC 2700
QY |||||
DB |||||
2704 CACCTGGACTTCGTTGATGAGCATCTCTGGTGTTCACCTCTCCCGCAGGATGGCTCAAT 2763
DB |||||
2701 CACCTGGACTTCGTTGATGAGCATCTCTGGTGTTCACCTCTCCCGCAGGATGGCTCAAT 2760
QY |||||
DB |||||
2764 GACTGTATGCAACAACGCGGAGTGTGGCAGCTGTGCTTTGCCATCCCGCGGGCCAC 2823
QY |||||
DB |||||
2761 GACTGTATGCAACAACGCGGAGTGTGGCAGCTGTGCTTTGCCATCCCGCGGGCCAC 2820
QY |||||
DB |||||
2824 CGCTGGCTGGCCTCTACACTACACCTTGACCCCGCAGCGCGCAACTGACAGCCGGCC 2883
DB |||||
2821 CGCTGGCTGGCCTCTACACTACACCTTGACCCCGCAGCGCGCAACTGACAGCCGGCC 2880
QY |||||
DB |||||
2884 ACCACTTCTTCTCTGTTCAGCCAGAAATCTGCCATCTGCTGATGATCCCGAGACACAG 2943
DB |||||
2881 ACCACTTCTTCTGTTCAGCCAGAAATCTGCCATCTGCTGATGATCCCGAGACACAG 2940
QY |||||
DB |||||
2944 CACAGCCGGATCTCATCTCTGCTGATGATGAGTGTGAGAACTCAAGCCATCGACTAT 3003
DB |||||
2941 CACAGCCGGATCTCATCTCTGCTGATGATGAGTGTGAGAACTCAAGCCATCGACTAT 3000
QY |||||
DB |||||
3004 GACCCACTGGACAGTTCTATCTACCTGGGTGATGGCGCCAGAACTCAAGCGAGCCAG 3063
DB |||||
3001 GACCCACTGGACAGTTCTATCTACCTGGGTGATGGCGCCAGAACTCAAGCGAGCCAG 3060
QY |||||
DB |||||
3064 GACGAGGGACCCAGCCCTTTGTTTGTGACCTCTTGAGCCAGGCGCAAAACCCAGACAGG 3123
DB |||||
3061 GACGAGGGACCCAGCCCTTTGTTTGTGACCTCTTGAGCCAGGCGCAAAACCCAGACAGG 3120
QY |||||
DB |||||
3124 CAGCCCGCAGACCTCAGCATCGACATCTACAGCGGGAACATCTGTTCTGAGAGTGGAGGCC 3183
DB |||||
3121 CAGCCCGCAGACCTCAGCATCGACATCTACAGCGGGAACATCTGTTCTGAGACGTGCGAGGCC 3180
QY |||||
DB |||||
3184 ACCAATACCATCAACGTTCCACAGGCTGAGCGGGGAAGCCATGGGGGTGTGCTGCTGGG 3243
DB |||||
3181 ACCAATACCATCAACGTTCCACAGGCTGAGCGGGGAAGCCATGGGGGTGTGCTGCTGGG 3240
QY |||||
DB |||||
3244 GACCGCGCAAGCCAGGGCCATCTGCTGTCACGCGGAGCGAGGTACTGTTACTTCAAC 3303
DB |||||
3241 GACCGCGCAAGCCAGGGCCATCTGCTGTCACGCGGAGCGAGGGTACTGTTACTTCAAC 3300
QY |||||
DB |||||
3304 AACATGAGGACCGGGCAGCCAAAGTTCGAAACGCGCAGCCCTGGAAGCCAGCGCGCGAG 3363
DB |||||
3301 AACATGAGGACCGGGCAGCCAAAGTTCGAAACGCGCAGCCCTGGAAGCCAGCGCGCGAG 3360
QY |||||
DB |||||
3364 GTCTCTTTTCAACACCGGCTCATCTCCGCTCTGTCGCTGTGGTGTGGACAAACACACTGGGC 3423
DB |||||
3361 GTCTCTTTTCAACACCGGCTCATCTCCGCTCTGTCGCTGTGGTGTGGTGTGGTGTGGG 3420
QY |||||
DB |||||
3424 AAGCTGTTCTGGTGTGGACGCGGACCTGAAGCGCATTTGAGAGTGTGACTGTGTCAGGGGCC 3483
DB |||||
3421 AAGCTGTTCTGGTGTGGACGCGGACCTGAAGCGCATTTGAGAGTGTGACTGTGTCAGGGGCC 3480
QY |||||
DB |||||
3484 AACCGCTTGACCTGGAGAGCGCAACATCTGTGACGCTCTGGGCTCAGCATCTGCTGGC 3543
DB |||||
3481 AACCGCTTGACCTGGAGAGCGCAACATCTGTGACGCTCTGGGCTCAGCATCTGCTGGC 3540

QY 3544 AAGCATCTTACTGATCGAACCGCCAGCAGAGATGATCGAGCGTGTGGAGAACACACC 3603
DB |||||
3541 AAGCATCTTACTGATCGAACCGCCAGCAGAGATGATCGAGCGTGTGGAGAACACACC 3600
QY |||||
DB |||||
3604 GGGGCAAGCGGACTCGCATCCAGGGCGTGTGGCCCACTCACTGGCATCCATCGAGTG 3663
DB |||||
3601 GGGGCAAGCGGACTCGCATCCAGGGCGTGTGGCCCACTCACTGGCATCCATCGAGTG 3660
QY |||||
DB |||||
3664 GAGGAAGTCAGCCTTGAGGAGTCTTCAGCCCAACCATGTGCCCTGTGACAAATGGTGGCTGC 3723
DB |||||
3661 GAGGAAGTCAGCCTTGAGGAGTCTTCAGCCCAACCATGTGCCCTGTGACAAATGGTGGCTGC 3720
QY |||||
DB |||||
3724 TCCCAATCTGTATTGCCAAGGGTGTGGGACACACAGGTGTCTATGGCCAGTGCACCTC 3783
DB |||||
3721 TCCCAATCTGTATTGCCAAGGGTGTGGGACACACAGGTGTCTATGGCCAGTGCACCTC 3780
QY |||||
DB |||||
3784 GTGCTCTTCAGAACTCTCTGACCTGTGAGAGCGCCACCTGTCTCCCGGACAGTTT 3843
DB |||||
3781 GTGCTCTTCAGAACTCTCTGACCTGTGAGAGCGCCACCTGTCTCCCGGACAGTTT 3840
QY |||||
DB |||||
3844 GATGTGCCACAGGGAGATCGACTGTATCCCGGGGCTGTGCGCTGTGACGGCTTTCC 3903
DB |||||
3841 GATGTGCCACAGGGAGATCGACTGTATCCCGGGGCTGTGCGCTGTGACGGCTTTCC 3900
QY |||||
DB |||||
3904 GAGTGCATGACAGAGCGAGAGGGCTGCCCGTGTGCTCCGCGCCAGTTCC 3963
DB |||||
3901 GAGTGCATGACAGAGCGAGAGGGCTGCCCGTGTGCTCCGCGCCAGTTCC 3960
QY |||||
DB |||||
3964 TGCGCGGGGTCTGAGTGTGTGACCTGCGCTGTGCGCTGCGAGCGGAGGACACTGTGAG 4023
DB |||||
3961 TGCGCGGGGTCTGAGTGTGTGACCTGCGCTGTGCGCTGCGAGCGGAGGACACTGTGAG 4020
QY |||||
DB |||||
4024 GACGCTCAGACGAGGGGACTGTGACGCGCATCTGCTGCCCAACCACTTCGGTGTGG 4083
DB |||||
4021 GACGCTCAGACGAGGGGACTGTGACGCGCATCTGCTGCCCAACCACTTCGGTGTGG 4080
QY |||||
DB |||||
4084 AGCGCCAGTGTCTCTCATCAACAGCAGTGTGCTCTTCCCGCATGTGTATCGACGCG 4143
DB |||||
4081 AGCGCCAGTGTCTCTCATCAACAGCAGTGTGCTCTTCCCGCATGTGTATCGACGCG 4140
QY |||||
DB |||||
4144 TCCGACGAGCTCATGTGTGAAATCACCAGCGGCTCTCAGACGACAGCGCGGCCACAGC 4203
DB |||||
4141 TCCGACGAGCTCATGTGTGAAATCACCAGCGGCTCTCAGACGACAGCGCGGCCACAGC 4200
QY |||||
DB |||||
4204 AGTGCATCGGGCGCTCATTTGGCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4263
DB |||||
4201 AGTGCATCGGGCGCTCATTTGGCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTAT 4260
QY |||||
DB |||||
4264 TTTGTGTCCAGCGCTGTGTGCGGCTATGCGGGGCAACGCGGCTCTTCCCGCAC 4323
DB |||||
4261 TTTGTGTCCAGCGCTGTGTGCGGCTATGCGGGGCAACGCGGCTCTTCCCGCAC 4320
QY |||||
DB |||||
4324 GAGTATGTCAGCGGACCCCGCATCGTGCCTCTCAATTTTCATAGCCCGGGGGTTCACG 4383
DB |||||
4321 GAGTATGTCAGCGGACCCCGCATCGTGCCTCTCAATTTTCATAGCCCGGGGGTTCACG 4380
QY |||||
DB |||||
4384 CATGGCCCTTTCAGAGCATCGCATGGGAAGTTCATGATGAGCTCGTGGAGCTGTG 4443
DB |||||
4381 CATGGCCCTTTCAGAGCATCGCATGGGAAGTTCATGATGAGCTCGTGGAGCTGTG 4440
QY |||||
DB |||||
4444 GGGGGCGGGGGGGGGTCCCTCTGTAGACCGGAACCACTGTCACAGGGGCTCTGTCACG 4503
DB |||||
4441 GGGGGCGGGGGGGGGTCCCTCTACGACCGGAACCACTGTCACAGGGGCTCTGTCACG 4500
QY |||||
DB |||||
4504 AGCTGTCTCAGCAAGAGGCGACGCTGTACCGCGGATTCCTGAAACCGCGCGCTTCCCG 4563
DB |||||
4501 AGCTGTCTCAGCAAGAGGCGACGCTGTACCGCGGATTCCTGAAACCGCGCGCTTCCCG 4560
QY |||||
DB |||||
4564 GCCACGAGCCCTCTCTGTCAACATGGACATGTTCTTACTCTTCAAACTTCGSGCCACT 4623
DB |||||
4561 GCCACGAGCCCTCTCTGTCAACATGGACATGTTCTTACTCTTCAAACTTCGSGCCACT 4620
QY |||||
DB |||||
4624 GCGAGACCGTACAGGCGCTCATCATTTGAGGAATGGCGCCCCCGACGACGCGCTTCGAGC 4683

QY 544 ATGTACTGGACAGACTGGGTGAGACGCCCGGATTGAGCGGGCAGAGGATGATGGACG 603
Db 541 ATGTACTGGACAGACTGGGTGAGACGCCCGGATTGAGCGGGCAGAGGATGATGGACG 600
QY 604 ACCCGAAGATCATTTGTGGACTCGGACATTTTACTTGGCCCAATGGACTGACCATGACCTG 663
Db 601 ACCCGAAGATCATTTGTGGACTCGGACATTTTACTTGGCCCAATGGACTGACCATGACCTG 660
QY 664 GAGGACGAAAGCTCTACTGGGCTGACGCCAAGCTCAGCTTCATCCACGTCGCCAACCTG 723
Db 661 GAGGACGAAAGCTCTACTGGGCTGACGCCAAGCTCAGCTTCATCCACGTCGCCAACCTG 720
QY 724 GACGGCTCTGTCGGCAGAAAGTGTGTGAGGAGAGCTTGACGACCCCTTCCGCTTGACG 783
Db 721 GACGGCTCTGTCGGCAGAAAGTGTGTGAGGAGAGCTTGACGACCCCTTCCGCTTGACG 780
QY 784 CTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCCATCCATGCGCTGCAAC 843
Db 781 CTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCCATCCATGCGCTGCAAC 840
QY 844 AAGCGCACTGGGGGAAAGAGAGATCTGTAGTGCCCTCTACTCAACCATGACATC 903
Db 841 AAGCGCACTGGGGGAAAGAGAGATCTGTAGTGCCCTCTACTCAACCATGACATC 900
QY 904 CAGGTGTGACGAGGAGCGGAGCCTTTCTTCCACACTCGTGTGAGGAGACAAATGCG 963
Db 901 CAGGTGTGACGAGGAGCGGAGCCTTTCTTCCACACTCGTGTGAGGAGACAAATGCG 960
QY 964 GCGTCTCCCACTGTGCTGTGTCGCCAAGCGAGCCTTTCTACACATGCGCTGCCCC 1023
Db 961 GCGTCTCCCACTGTGCTGTGTCGCCAAGCGAGCCTTTCTACATGCGCTGCCCC 1020
QY 1024 ACGGTGTGCACTGACGACAAACGCGAGGACGTGTAAAGCAGGAGCGCGAGAGTGCTG 1083
Db 1021 ACGGTGTGCACTGACGACAAACGCGAGGACGTGTAAAGCAGGAGCGCGAGAGTGCTG 1080
QY 1084 CTGCTGGCCCGGCGACGGAACCTACGAGGATCTCGCTGGAACCGCGGACTTACCGAC 1143
Db 1081 CTGCTGGCCCGGCGACGGAACCTACGAGGATCTCGCTGGAACCGCGGACTTACCGAC 1140
QY 1144 ATCGTGTGCACTGACGACATCCGCGACGCAATGGCCATCGACTAGCAGCCGCTAGAG 1203
Db 1141 ATCGTGTGCACTGACGACATCCGCGACGCAATGGCCATCGACTAGCAGCCGCTAGAG 1200
QY 1204 GCGTATGTCTACTGACAGATGACGAGGTGCGGCGCATCCGACGCGGTACTGGAACGGG 1263
Db 1201 GCGTATGTCTACTGACAGATGACGAGGTGCGGCGCATCCGACGCGGTACTGGAACGGG 1260
QY 1264 TCTGGGGCGCAGACGCTGGTCAACCCGAGATCAACGACCCCGATGCGGTGCGAC 1323
Db 1261 TCTGGGGCGCAGACGCTGGTCAACCCGAGATCAACGACCCCGATGCGGTGCGAC 1320
QY 1324 TGGGTGGCCGAAACCTCTACTGGACCGACCGGACCGGACCGCATCGAGGTGACGCGC 1383
Db 1321 TGGGTGGCCGAAACCTCTACTGGACCGGACCGGACCGGACCGCATCGAGGTGACGCGC 1380
QY 1384 CTCAACGGCACCTCCGCAAGATCTGTGTGTCGAGGACCTGGAACGACCCCGACCTC 1443
Db 1381 CTCAACGGCACCTCCGCAAGATCTGTGTGTCGAGGACCTGGAACGACCCCGACCTC 1440
QY 1444 GCACTGCAACCCGCTGATGGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAATC 1503
Db 1441 GCACTGCAACCCGCTGATGGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAATC 1500
QY 1504 GAGTGTGCCAACTTGGATGGGAGGAGCGGCTGTGTGCTGCTCAATGCGCTCCCTCGGGTGG 1563
Db 1501 GAGTGTGCCAACTTGGATGGGAGGAGCGGCTGTGTGCTGCTCAATGCGCTCCCTCGGGTGG 1560
QY 1564 CCGAACGGCCTGGCCCTGGAACCTGACGAGGAGGAGCTTACTTGGGAGACGCCAAGACA 1623
Db 1561 CCGAACGGCCTGGCCCTGGAACCTGACGAGGAGGAGCTTACTTGGGAGACGCCAAGACA 1620

QY 1624 GACAGATCGAGCTGATCAATGTTGATGGGACGAGGCGGACCTCTCTGGAGACAAG 1683
Db 1621 GACAGATCGAGCTGATCAATGTTGATGGGACGAGGCGGACCTCTCTGGAGACAAG 1680
QY 1684 CTCCCGCACATTTTCGGGTTTCACTGCTGGGGACTTTCATCTACTGAGCTGACTGGCAG 1743
Db 1681 CTCCCGCACATTTTCGGGTTTCACTGCTGGGGACTTTCATCTACTGAGCTGACTGGCAG 1740
QY 1744 CGCCGACGACTCGAGCGGGTGCAAAAGGTCAAGGCCACCGCGGAGCTCATCATTTGACCAAG 1803
Db 1741 CGCCGACGACTCGAGCGGGTGCAAAAGGTCAAGGCCACCGCGGAGCTCATCATTTGACCAAG 1800
QY 1804 CTGCCCGACCTGATGGGCTCAAAAGCTGTGAATGTGGCAAGGCTGTCGGAACCAACCCG 1863
Db 1801 CTGCCCGACCTGATGGGCTCAAAAGCTGTGAATGTGGCAAGGCTGTCGGAACCAACCCG 1860
QY 1864 TGTGCGGACAGAAACGGGGGTGAGCCACCTGTGTCTTTTTCACACCCCAACCCCG 1923
Db 1861 TGTGCGGACAGAAACGGGGGTGAGCCACCTGTGTCTTTTTCACACCCCAACCCCG 1920
QY 1924 TGTGGCTGCCCATCGGCTGGAGCTGTGTGATGACATGAAGACCTTGATCGTCTGAG 1983
Db 1921 TGTGGCTGCCCATCGGCTGGAGCTGTGTGATGACATGAAGACCTTGATCGTCTGAG 1980
QY 1984 GCTTCTGTGTCTTACCCAGCAGAGCGCCATCCACAGGATCTCCCTCGAGACCAATAAC 2043
Db 1981 GCTTCTGTGTCTTACCCAGCAGAGCGCCATCCACAGGATCTCCCTCGAGACCAATAAC 2040
QY 2044 AACGACGTGGCATCCGCTCACGGGCGTCAAGGAGGCTCAGCCCTGGACTTTGATGTG 2103
Db 2041 AACGACGTGGCATCCGCTCACGGGCGTCAAGGAGGCTCAGCCCTGGACTTTGATGTG 2100
QY 2104 TCCAAACACACATCTACTGACAGACGTCAGCTGTGAAGACCATCAGCCGCGCTTCATG 2163
Db 2101 TCCAAACACACATCTACTGACAGACGTCAGCTGTGAAGACCATCAGCCGCGCTTCATG 2160
QY 2164 AACGGAGCTCGGTGGAGACGTCGTGTGAGTTTGGCCCTTGACTACCCGAGGGATGCCC 2223
Db 2161 AACGGAGCTCGGTGGAGACGTCGTGTGAGTTTGGCCCTTGACTACCCGAGGGATGCCC 2220
QY 2224 GTTGAATGGATGGCAAGAACCTCTACTGAGCGGACACTGGGACCAACAGAAATCGAATG 2283
Db 2221 GTTGAATGGATGGCAAGAACCTCTACTGAGCGGACACTGGGACCAACAGAAATCGAATG 2280
QY 2284 GCGCGCTGGAGCGGACGTCGCGCAAGTCTCGTGTGGAGGACTTGGACAAACCCGAGG 2343
Db 2281 GCGCGCTGGAGCGGACGTCGCGCAAGTCTCGTGTGGAGGACTTGGACAAACCCGAGG 2340
QY 2344 TCGTGGCCCTGGATCCCAACAGGGCTACATCTACTGGACCGAGTGGGGCGGCAAGCCG 2403
Db 2341 TCGTGGCCCTGGATCCCAACAGGGCTACATCTACTGGACCGAGTGGGGCGGCAAGCCG 2400
QY 2404 AGGATCGTGGCGCTTCATGGAACGGGACCAACTGTCATGAGCTGTGGTGGACAGGTGGGC 2463
Db 2401 AGGATCGTGGCGCTTCATGGAACGGGACCAACTGTCATGAGCTGTGGTGGACAGGTGGGC 2460
QY 2464 CGGGCAACGACCTTCAACATTTGACTAGCTGACAGCGCTCTACTGACCGACCTGGAC 2523
Db 2461 CGGGCAACGACCTTCAACATTTGACTAGCTGACAGCGCTCTACTGACCGACCTGGAC 2520
QY 2524 ACCAATGATGAGTCTGCTCCAAACATGTGGGTGAGGAGCGGGTCTGTGATTTGCCACGAT 2583
Db 2521 ACCAATGATGAGTCTGCTCCAAACATGTGGGTGAGGAGCGGGTCTGTGATTTGCCACGAT 2580
QY 2584 CTCCGCGACCGGCTGCTGACCGGAGTACAGGATTAATCTACTGACAGACTGGAT 2643
Db 2581 CTCCGCGACCGGCTGCTGACCGGAGTACAGGATTAATCTACTGACAGACTGGAT 2640
QY 2644 CTGACAGCATTTGAGCGGGCGCAAGAGCTAGCGGCCGGAACCGACCTCATCCAGGGC 2703
Db 2641 CTGACAGCATTTGAGCGGGCGCAAGAGCTAGCGGCCGGAACCGACCTCATCCAGGGC 2700
QY 2704 CACCTGGACTTCTGTGATGAGCATCTCTGTGTGTTCCAATCTCTCCCGCAGAGATGGCCTCAAT 2763

[illegible]

3781	GTGTCCTCTCGAGAACTGCTGAC	CTGTGGAGAGCGCGCCACCTGCTCCCGGACCAAGTTT	3843
3844	GCATGTGCCACAGGGAGATCGA	CTGTATCCCGGGGCTTGGCGCTGTGACGGCTTTCCC	3903
3841	GCATGTGCCACAGGGAGATCGA	CTGTATCCCGGGGCTTGGCGCTGTGACGGCTTTCCC	3900
3904	GAGTGCATGATCAGACGACGAG	AGGAGGGCTGCCCGTGTGCTTCGGCGGCCAGTTCCCC	3963
3901	GAGTGCATGATCAGACGACGAG	AGGAGGGCTGCCCGTGTGCTTCGGCGGCCAGTTCCCC	3960
3964	TGCGCGCGGGGTGAGTGTGTGG	ACCTGCGCTGCGCTGCACGCGCAGGAGCAGACTGTTCAG	4023
3961	TGCGCGCGGGGTGAGTGTGTGG	ACCTGCGCTGCGCTGCACGCGCAGGAGCAGACTGTTCAG	4020
4024	GACCGCTCAGACGAGCGGACCT	GTGACGCCATCTTGCTGCGCCAAACCAAGTTCCGGTGTGCG	4083
4021	GACCGCTCAGACGAGCGGACCT	GTGACGCCATCTTGCTGCGCCAAACCAAGTTCCGGTGTGCG	4080
4084	AGCGGCCAGTGTCTTCATCAAA	CACAGCAGTGGACCTCTTCCCGAGCTGTATCGACGGC	4143
4081	AGCGGCCAGTGTCTTCATCAAA	CACAGCAGTGGAGCTCTTCCCGAGCTGTATCGACGGC	4140
4144	TCCGACAGCTCATGTGTGAAAT	CACCAGCGCCCTCAGACGACGACCGCGGCCACAGC	4203
4141	TCCGACAGCTCATGTGTGAAAT	CACCAGCGCCCTCAGACGACGACCGCGGCCACAGC	4200
4204	AGTGCCATCGGGGCCGTGATGG	GCATCATCTCTCTCTTCGTGATGGGTGGTGTCTAT	4263
4201	AGTGCCATCGGGGCCGTGATGG	GCATCATCTCTCTCTTCGTGATGGGTGGTGTCTAT	4260
4264	TTTGTGTGCAGCGCGTGTGTGC	AGCGCTATGCGGGGCCAAACGGGCGCTTCCCGCAC	4323
4261	TTTGTGTGCAGCGCGTGTGTGC	AGCGCTATGCGGGGCCAAACGGGCGCTTCCCGCAC	4320
4324	GAGTATGTACGGGACCCCGCAG	CGTGCCTCAATTTTCATAGCCCCGGGGGTGCCAG	4383
4321	GAGTATGTACGGGACCCCGCAG	CGTGCCTCAATTTTCATAGCCCCGGGGGTGCCAG	4380
4384	CATGGCCCCCTTACAGGCATCG	ATGCGGAAAGTCCATGATGAGCTCCGTGAGCTGATG	4443
4381	CATGGCCCCCTTACAGGCATCG	ATGCGGAAAGTCCATGATGAGCTCCGTGAGCTGATG	4440
4444	GGGGGCGGGCGGGGTGCCCTGT	ACGACCGAAACACGTCAAGGGGCTCGTCCAGC	4503
4441	GGGGGCGGGCGGGGTGCCCTGT	ACGACCGAAACACGTCAAGGGGCTCGTCCAGC	4500
4504	AGCTCGTCCAGCAGAGGCCAG	CGCTGTACCCCGCGATCTGTAAACCGCGCCCTCCCG	4563
4501	AGCTCGTCCAGCAGAGGCCAG	CGCTGTACCCCGCGATCTGTAAACCGCGCCCTCCCG	4560
4564	GCCACGACCCCTCCCTGTAAAC	ATGGAATGTTTACTCTTCAACATTCGGGCCACT	4623
4561	GCCACGACCCCTCCCTGTAAAC	ATGGAATGTTTACTCTTCAACATTCGGGCCACT	4620
4624	GCGAGACCGTACAGGCCCTTAC	ATTCGAGGAATGCGCCCCCGACGACGCCCTGCAGC	4683
4621	GTGAGACCGTACAGGCCCTTAC	ATTCGAGGAATGCGCCCCCGACGACGCCCTGCAGC	4680
4684	ACCGAGCTGTGTACAGCGACTA	CACGGCCAGCGCTGTGGAAGGCCAGCAAGTACTACCTG	4743
4681	ACCGAGCTGTGTACAGCGACTA	CACGGCCAGCGCTGTGGAAGGCCAGCAAGTACTACCTG	4740
4744	GATTTGAACTCGGACTCAGACCC	CTATTCACCCCAACCGCCCCACAGCCAGTACTCTG	4803
4741	GATTTGAACTCGGACTCAGACCC	CTATTCACCCCAACCGCCCCACAGCCAGTACTCTG	4800
4804	TCGGCGGAGGACAGCTGCGCC	CGCTTCCCGCCACCGAGAGGAGCTACTTCCATCTCTTC	4863
4801	TCGGCGGAGGACAGCTGCGCC	CGCTTCCCGCCACCGAGAGGAGCTACTTCCATCTCTTC	4860
4864	CCGCCCCCTCGTCCCTCGACGG	ACTCATCTGTGACCTCGCGCGGGCCACTCTGGGTTTC	4923
4861	CCGCCCCCTCGTCCCTCGACGG	ACTCATCTGTGACCTCGCGCGGGCCACTCTGGGTTTC	4920

Db 1561 CCCAACGGCCTGGCCCTGGACCTCGCAGAGGGGAGCTCTACTGGGGAGACGCCAAGACA 1620
Qy 1624 GACAAAGATCGAGGTGATCAATGTTGATCGGACGAGAGCGGAGCCCTCTGGAGACAAAG 1683
Db 1621 GACAAAGATCGAGGTGATCAATGTTGATGGGACGAGAGCGGAGCCCTCTGGAGGACAAAG 1680
Qy 1684 CTCGCGCACATTTTCGGGTTTCACTGCTGCTGGGGGACTTACTACTGAGACTGACTGGCAG 1743
Db 1681 CTCGCGCACATTTTCGGGTTTCACTGCTGCTGGGGGACTTACTACTGAGACTGACTGGCAG 1740
Qy 1744 CGCGCAGCATCGAGCGGGTGCACAAAGGTCAAGGCTCAAGCGCAGCGGAGCTGATCATTTGACCAAG 1803
Db 1741 CGCGCAGCATCGAGCGGGTGCACAAAGGTCAAGGCTCAAGCGCAGCGGAGCTGATCATTTGACCAAG 1800
Qy 1804 CTGCCCGACCTGATGGGGCTCAAGCTGTGAATGTGGCCAAAGGTCTCGGAAACCAACCCG 1863
Db 1801 CTGCCCGACCTGATGGGGCTCAAGCTGTGAATGTGGCCAAAGGTCTCGGAAACCAACCCG 1860
Qy 1864 TGTGCGGACAGAAACGGGGGTGACGCCAACCCTGTGCTTCTTTCACACCCCAACCGAACCCGG 1923
Db 1861 TGTGCGGACAGAAACGGGGGTGACGCCAACCCTGTGCTTCTTTCACACCCCAACCGAACCCGG 1920
Qy 1924 TGTGGCTGCCCCATCGGCTCGAGCTGTGAGTGTGATGACATGAAGACCTGATCGTGCTGAG 1983
Db 1921 TGTGGCTGCCCCATCGGCTCGAGCTGTGAGTGTGATGACATGAAGACCTGATCGTGCTGAG 1980
Qy 1984 GCCTTCTTGGTCTTCCACAGCAGAGCGGCCATCCACAGGATCTCCCTCGAGACCAATAAC 2043
Db 1981 GCCTTCTTGGTCTTCCACAGCAGAGCGGCCATCCACAGGATCTCCCTCGAGACCAATAAC 2040
Qy 2044 AACGACGTGGCCATCCCGCTCAACGGGCTCAAGGAGGCTCAGCCCTGGACTTTGATGTG 2103
Db 2041 AACGACGT -GCCATCCCGCTCAACGGGCTCAAGGAGGCTCAGCCCTGGACTTTGATGTG 2099
Qy 2104 TCCAAACACACATCTATGACAGACGTGAGCTGAAAGACCATCAGCGCGGCTTCATG 2163
Db 2100 TCCAAACACACATCTATGACAGACGTGAGCTGAAAGACCATCAGCGCGGCTTCATG 2159
Qy 2164 AACGGGAGCTGGTGGAGCAGTGTGAGTTGGCTTTGACTACCCGAGGGGATGGCC 2223
Db 2160 AACGGGAGCTGGTGGAGCAGTGTGAGTTGGCTTTGACTACCCGAGGGGATGGCC 2219
Qy 2224 GTTGACTGGATGGGCAAGAACTCTACTGGGCGCACACTGGGACCAACAGAAATCGAAGTG 2283
Db 2220 GTTGACTGGATGGGCAAGAACTCTACTGGGCGCACACTGGGACCAACAGAAATCGAAGTG 2279
Qy 2284 GCGGGCTGGAACGGGCAAGTCCGGCAAGTCTCGTGTGGAGGAGCTTGGACAAACCCGAGG 2343
Db 2280 GCGGGCTGGAACGGGCAAGTCCGGCAAGTCTCGTGTGGAGGAGCTTGGACAAACCCGAGG 2339
Qy 2344 TCGCTGGCCCTGGATCCCAACAGGGCTACATCTACTGGACCGGAGTGGGGCGGCAAGCCG 2403
Db 2340 TCGCTGGCCCTGGATCCCAACAGGGCTACATCTACTGGACCGGAGTGGGGCGGCAAGCCG 2399
Qy 2404 AGGATCGTGGGGCTTTCATGGAACGGGACCAACTGATGAGCGCTGGTGGAACAAAGTGGGC 2463
Db 2400 AGGATCGTGGGGCTTTCATGGAACGGGACCAACTGATGAGCGCTGGTGGAACAAAGTGGGC 2459
Qy 2464 CGGGCAACAGACCTCAACATTTGACTACGCTGACCGAGTGTGAGTGGAGTGGGGCGGCAAGCCG 2523
Db 2460 CGGGCAACAGACCTCAACATTTGACTACGCTGACCGAGCGCTTACTTGGACCGACCTGGAC 2519
Qy 2524 ACCAATCATGATCGATCGTCCAAATGCTGGGTGAGGAGCGGGTCTGATTTGCCGACGAT 2583
Db 2520 ACCAATCATGATCGATCGTCCAAATGCTGGGTGAGGAGCGGGTCTGATTTGCCGACGAT 2579
Qy 2584 CTCGCGCACCCGTTCCGCTGTGACGAGTACAGCGATTTATATCTACTGACACAGACTGGAAAT 2643
Db 2580 CTCGCGCACCCGTTCCGCTGTGACGAGTACAGCGATTTATATCTACTGACACAGACTGGAAAT 2639
Qy 2644 CTGCAACAGCATTTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCCCTCATCCAGGGC 2703
Db 2640 CTGCAACAGCATTTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCCCTCATCCAGGGC 2699

Qy 2704 CACCTGGAATTCGTGATGGACATCCTGTGTGTTTCCATCTCTCCCGCCAGGATGGCTCAAT 2763
Db 2700 CACCTGGAATTCGTGATGGACATCCTGTGTGTTTCCATCTCTCCCGCCAGGATGGCTCAAT 2759
Qy 2764 GACTGTATGACAAACAAACGGGCAAGTGTGGGAGCTGTGGCTTGGCATCCCGCGGGCCAC 2823
Db 2760 GACTGTATGACAAACAAACGGGCAAGTGTGGGAGCTGTGGCTTGGCATCCCGCGGGCCAC 2819
Qy 2824 CGCTGCGGCTGCGCTCACAACCTGAGACCCCGCCAGCGCCCAACCTGAGCGAGCCGCCC 2883
Db 2820 CGCTGCGGCTGCGCTCACAACCTGAGACCCCGCCAGCGCCCAACCTGAGCGAGCCGCCC 2879
Qy 2884 ACCACCTTCTTGTCTGTTTTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGGAGACACAG 2943
Db 2880 ACCACCTTCTTGTCTGTTTTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGGAGACACAG 2939
Qy 2944 CACAGCCGGATCTCATCTCTGCTGCTGAGTGAATGAGAAACGTTCAAGCCATCGATAT 3003
Db 2940 CACAGCCGGATCTCATCTCTGCTGCTGAGTGAATGAGAAACGTTCAAGCCATCGATAT 2999
Qy 3004 GACCACTGAGACAAAGTTCATCTACTGAGTGGATGGGCGCCAGAAACATCAAGCGAGCCAAAG 3063
Db 3000 GACCACTGAGACAAAGTTCATCTACTGAGTGGATGGGCGCCAGAAACATCAAGCGAGCCAAAG 3059
Qy 3064 GACGACGGGACCCAGCCCTTGTGTTTTCAGCTCTCTGAGCCAAAGCCAAACCCAGACAGG 3123
Db 3060 GACGACGGGACCCAGCCCTTGTGTTTTCAGCTCTCTGAGCCAAAGCCAAACCCAGACAGG 3119
Qy 3124 CAGCCCAACGACCTCAGCATCGACATCTACAGCCGGAACACTGTTCTGGACGCTGGAGGCC 3183
Db 3120 CAGCCCAACGACCTCAGCATCGACATCTACAGCCGGAACACTGTTCTGGACGCTGGAGGCC 3179
Qy 3184 ACCAATACCATCAAGCTTCCACAGCTGAGCGGGGAGCCATGGGGGTGGTCTGGTGGG 3243
Db 3180 ACCAATACCATCAAGCTTCCACAGCTGAGCGGGGAGCCATGGGGGTGGTCTGGTGGG 3239
Qy 3244 GACCGCACAAAGCCAGGGCCATCGTCTGTAACCGCGAGCGAGGATCTCTGATCTTACCC 3303
Db 3240 GACCGCACAAAGCCAGGGCCATCGTCTGTAACCGCGAGCGAGGATCTCTGATCTTACCC 3299
Qy 3304 AACATGAGGACCGGGCAGCAAGATCGAAGCGCGAGCTTGGACCGGACCGAGCGGAG 3363
Db 3300 AACATGAGGACCGGGCAGCAAGATCGAAGCGCGAGCTTGGACCGGACCGAGCGGAG 3359
Qy 3364 GTCTCTTTCACCGCGGCTCATCCGCGCTGTGGCTGTGGTGGGACCAACACTGGGC 3423
Db 3360 GTCTCTTTCACCGCGGCTCATCCGCGCTGTGGCTGTGGTGGGACCAACACTGGGC 3419
Qy 3424 AAGCTGTTCTGGGTGGAGCGGACCTGAAGCGCATTTGAGAGCTGTGACTGTGAGGGGCC 3483
Db 3420 AAGCTGTTCTGGGTGGAGCGGACCTGAAGCGCATTTGAGAGCTGTGACTGTGAGGGGCC 3479
Qy 3484 AACCGCTGACCTTGGAGGACCGCAACATCGTGCAGCTCTTGGGCTTGAACATCTTGGC 3543
Db 3480 AACCGCTGACCTTGGAGGACCGCAACATCGTGCAGCTCTTGGGCTTGAACATCTTGGC 3539
Qy 3544 AAGCATCTCTACTGATTCGACCGCAGCAGCAGATGATTCGAGCGCTGTGGAGAAACACCC 3603
Db 3540 AAGCATCTCTACTGATTCGACCGCAGCAGCAGATGATTCGAGCGCTGTGGAGAAACACCC 3599
Qy 3604 GGGGACAAAGCGGACTCGCATCCAGGGCCGTGTGGCCCAACCTCACTGGCATCCATGCACTG 3663
Db 3600 GGGGACAAAGCGGACTCGCATCCAGGGCCGTGTGGCCCAACCTCACTGGCATCCATGCACTG 3659
Qy 3664 GAGGAAGTTCAGCTTGGAGGAGTCTCAGCCCAACCCATGTGCCGCTGAACAATGTTGGCTGC 3723
Db 3660 GAGGAAGTTCAGCTTGGAGGAGTCTCAGCCCAACCCATGTGCCGCTGAACAATGTTGGCTGC 3719
Qy 3724 TCCCAACATCTGATTTGCCAAGGGTGTGAGGACACCAACGCTGCTCATGCCCCAGTCCACCTC 3783
Db 3720 TCCCAACATCTGATTTGCCAAGGGTGTGAGGACACCAACGCTGCTCATGCCCCAGTCCACCTC 3779

QY 484 CTCCTTCTGGCAGGACCTTGACAGCCTTAGGGCCATCGCCTTGAGACCCCGCTCACGGGTAC 543
DB 481 CTCCTTCTGGCAGGACCTTGACAGCCTTAGGGCCATCGCCTTGAGACCCCGCTCACGGGTAC 540
QY 544 ATGTACTGGACAGACTGGGGTGGAGCGCCCGGATTTAGCGGGCAGGAGTGGATGGCAGC 603
DB 541 ATGTACTGGACAGACTGGGGTGGAGCGCCCGGATTTAGCGGGCAGGAGTGGATGGCAGC 600
QY 604 ACCCGGAGAGATCATTTGTCGACCTCGACATTTTACTGGCCCAATGGACTGACCATCGACCTG 663
DB 601 ACCCGGAGAGATCATTTGTCGACCTCGACATTTTACTGGCCCAATGGACTGACCATCGACCTG 660
QY 664 GAGGAGCAGAACTCTACTTGGGCTGACGCGCAAGCTCAGCTTCATCCACGCTGCCAACCTG 723
DB 661 GAGGAGCAGAACTCTACTTGGGCTGACGCGCAAGCTCAGCTTCATCCACGCTGCCAACCTG 720
QY 724 GACGGCTCTGTTCCGCGAAGAGTGTGGAGGCAAGCTTGACGCAAGCTTGGCCCTGACG 783
DB 721 GACGGCTCTGTTCCGCGAAGAGTGTGGAGGCAAGCTTGACGCAAGCTTGGCCCTGACG 780
QY 784 CTCCTCGGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCCATCGCTGCAAC 843
DB 781 CTCCTCGGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCCATCGCTGCAAC 840
QY 844 AAGCGCACTGGGGGGAAGAGGAGAGATCTCTGAGTGCCCTCTACTCACCCATGGACATC 903
DB 841 AAGCGCACTGGGGGGAAGAGGAGAGATCTCTGAGTGCCCTCTACTCACCCATGGACATC 900
QY 904 CAGGTGTGAGCAGGAGCGGAGCCTTTCTTCCACACTCGCTGTGGAGGAGCAATGGC 963
DB 901 CAGGTGTGAGCAGGAGCGGAGCCTTTCTTCCACACTCGCTGTGGAGGAGCAATGGC 960
QY 964 GGCTGCTCCACCTGTGCTGCTGCTGCCCAAGCGAGCCTTTCTACACATGCGCTGCCCC 1023
DB 961 GGCTGCTCCACCTGTGCTGCTGCTGCCCAAGCGAGCCTTTCTACACATGCGCTGCCCC 1020
QY 1024 ACGGCTGTGACAGCTGCAGACAAACGGCAGGAGCTGTAAAGCAGGAGCCGAGGAGTGCTG 1083
DB 1021 ACGGCTGTGACAGCTGCAGACAAACGGCAGGAGCTGTAAAGCAGGAGCCGAGGAGTGCTG 1080
QY 1084 CTGCTGGCCCGCGCGAGCGGACCTTACGAGGAGTCTCGCTTGGACACGCCCGGACTTACCGAC 1143
DB 1081 CTGCTGGCCCGCGCGAGCGGACCTTACGAGGAGTCTCGCTTGGACACGCCCGGACTTACCGAC 1140
QY 1144 ATCGTGTGACAGTGGACAGCATCGGCAAGCCTTGGCCATCGACTACGACCGCTAGAG 1203
DB 1141 ATCGTGTGACAGTGGACAGCATCGGCAAGCCTTGGCCATCGACTACGACCGCTAGAG 1200
QY 1204 GGCTATGTCTACTGGACAGATGACGAGGTGGGGCCATCCGCGAGGCGGTACTTGGACGGG 1263
DB 1201 GGCTATGTCTACTGGACAGATGACGAGGTGGGGCCATCCGCGAGGCGGTACTTGGACGGG 1260
QY 1264 TCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGTGCAC 1323
DB 1261 TCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGTGCAC 1320
QY 1324 TGGGTGGCCGCAAACTCTACTGGACCGACACGGGCAAGGACCGCATCGAGTGGACGCG 1383
DB 1321 TGGGTGGCCGCAAACTCTACTGGACCGACACGGGCAAGGACCGCATCGAGTGGACGCG 1380
QY 1384 CTCACAGGCACTCCCGCAAGATCTCTGTGTGGAGGACCTGGACAGCCCGAGCCCATC 1443
DB 1381 CTCACAGGCACTCCCGCAAGATCTCTGTGTGGAGGACCTGGACAGCCCGAGCCCATC 1440
QY 1444 GCATGCAACCCCGTATGGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAAATC 1503
DB 1441 GCATGCAACCCCGTATGGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAAATC 1500
QY 1504 GAGTGTGCCAATCTTGATGGGCGAGAGCGGCGTGTGCTGCTCAATGCTCTCGGTGG 1563
DB 1501 GAGTGTGCCAATCTTGATGGGCGAGAGCGGCGTGTGCTGCTCAATGCTCTCGGTGG 1560

QY 1564 CCCAACGGCCTGGCCCTGGACCTTGACGAGGGGAGCTTACTTGGGAGAGCCCAAGACA 1623
DB 1561 CCCAACGGCCTGGCCCTGGACCTTGACGAGGGGAGGAGCTTACTTGGGAGAGCCCAAGACA 1620
QY 1624 GACAGATTCGAGGTGATCAATGTTGATGGGACGAAAGAGCGGAGCCCTCTCTGGAGGACAAG 1683
DB 1621 GACAGATTCGAGGTGATCAATGTTGATGGGACGAAAGAGCGGAGCCCTCTCTGGAGGACAAG 1680
QY 1684 CTCCTCGCACATTTTTCGGGTTCAACGCTGCTGGGGGACTTCTACTTGTGACTGACTGGCAG 1743
DB 1681 CTCCTCGCACATTTTTCGGGTTCAACGCTGCTGGGGGACTTCTACTTGTGACTGACTGGCAG 1740
QY 1744 CGCCGACGATCGAGCGGGTGCACAAAGTCAAGGCCAGCCGCGGACGTTCATCTATTGACCAG 1803
DB 1741 CGCCGACGATCGAGCGGGTGCACAAAGTCAAGGCCAGCCGCGGACGTTCATCTATTGACCAG 1800
QY 1804 CTGCCCAGACTGATGGGCTCAAAAGCTGTGTAATGTGGCCAAAGGTCTGTCGGAACCAACCCG 1863
DB 1801 CTGCCCAGACTGATGGGCTCAAAAGCTGTGTAATGTGGCCAAAGGTCTGTCGGAACCAACCCG 1860
QY 1864 TGTGCGGACAGAAACGGGGGTGAGCCACCTGTGCTTCTTTCACACCCACGCAACCCGG 1923
DB 1861 TGTGCGGACAGAAACGGGGGTGAGCCACCTGTGCTTCTTTCACACCCACGCAACCCGG 1920
QY 1924 TGTGGCTGCCCATCGGCTGGAGCTGCTGAGTGAATGATGAAGACCTGTCATCTGCTGAG 1983
DB 1921 TGTGGCTGCCCATCGGCTGGAGCTGCTGAGTGAATGATGAAGACCTGTCATCTGCTGAG 1980
QY 1984 GCTTCTTGGTCTTCAACGAGAGCGCCATTCACAGGATCTCTCTCGAGACCAATTAAC 2043
DB 1981 GCTTCTTGGTCTTCAACGAGAGCGCCATTCACAGGATCTCTCTCGAGACCAATTAAC 2040
QY 2044 AACGAGCTGGCCATCCGCTCACCGGGCTCAAGGAGGCTCAGCCCTCGACTTTTCATGTG 2103
DB 2041 AACGAGCTGGCCATCCGCTCACCGGGCTCAAGGAGGCTCAGCCCTCGACTTTTCATGTG 2099
QY 2104 TCCAAACCAACATCTACTTGGACAGAGCTCAGCCCTGAAAGACCATCAGCGCGGCTTCATG 2163
DB 2100 TCCAAACCAACATCTACTTGGACAGAGCTCAGCCCTGAAAGACCATCAGCGCGGCTTCATG 2159
QY 2164 AACGGAGCTCGGTGGAGACAGTGTGGAGTTTGGCTTGTGACTACCCCGAGGGCATGGCC 2223
DB 2160 AACGGAGCTCGGTGGAGACAGTGTGGAGTTTGGCTTGTGACTACCCCGAGGGCATGGCC 2219
QY 2224 GTTGACTGAGTGGGCAAGAACCTCTACTTGGCCGACACTGGGACCAACAGCAATCGAAGTG 2283
DB 2220 GTTGACTGAGTGGGCAAGAACCTCTACTTGGCCGACACTGGGACCAACAGCAATCGAAGTG 2279
QY 2284 GCGGGCTGGACGGGCGAGTTCCGGCAAGTCTCGTGTGGAGGAGCTTGGACAAACCCGAGG 2343
DB 2280 GCGGGCTGGACGGGCGAGTTCCGGCAAGTCTCGTGTGGAGGAGCTTGGACAAACCCGAGG 2339
QY 2344 TCGCTGGCCCTGGATTCACCAACGAGGCTACATCTACTGGACCGAGTGGGGCGGCAAGCGG 2403
DB 2340 TCGCTGGCCCTGGATTCACCAACGAGGCTACATCTACTTGGACCGAGTGGGGCGGCAAGCGG 2399
QY 2404 AGGATCGTGGGGCTTCATGAGCGGGACCAACTGCTGATGACGCTGGTGCACAAAGTGGGC 2463
DB 2400 AGGATCGTGGGGCTTCATGAGCGGGACCAACTGCTGATGACGCTGGTGCACAAAGTGGGC 2459
QY 2464 CGGGCCAAACGACCTCACCATTTGACTACGCTGACCGAGCCCTCTACTGGACCGACCTGGAC 2523
DB 2460 CGGGCCAAACGACCTCACCATTTGACTACGCTGACCGAGCCCTCTACTGGACCGACCTGGAC 2519
QY 2524 ACCAAACATGATCGAGTCTGCAACATGCTGGGTTCAGGAGCGGGTCTGATTTGCCAGCAT 2583
DB 2520 ACCAAACATGATCGAGTCTGCAACATGCTGGGTTCAGGAGCGGGTCTGATTTGCCAGCAT 2579
QY 2584 CTCCTCGCACCCGTTCTGGTCTGACGAGTACAGGATTTATATCTACTTGGACAGACTGGAAT 2643
DB 2580 CTCCTCGCACCCGTTCTGGTCTGACGAGTACAGGATTTATATCTACTTGGACAGACTGGAAT 2639
QY 2644 CTGACACAGCAATTGAGCGGGCCGACAAAGACTAGCGGCGGAAACCGCACCTCTCATCCAGGCG 2703


```
QY 4864 CCGCCCCCTCGTCCCTCGACGGACTCATCTGACCTCGCGCGGGCCACTCTGGCTTC 4923
Db 4859 CCGCCCCCTCGTCCCTCGACGGACTCATCTGACCTCGCGCGGGCCACTCTGGCTTC 4918
QY 4924 TCTGTGCCCTGTAATAAGTATTTAAATATGAACAAAGAAAAAATATATTTATGATTTA 4983
Db 4919 TCTGTGCCCTGTAATAAGTATTTAAATATGAACAAAGAAAAAATATATTTATGATTTA 4978
QY 4984 AAAATAAATAATATGGGATTTTAAAAACATGAGAAATGGAATGTGATGGGTGGGC 5043
Db 4979 AAAATAAATAATATGGGATTTTAAAAACATGAGAAATGGAATGTGATGGGTGGGC 5038
QY 5044 AGGGCTGGGAGAACTTTGTA 5063
Db 5039 AGGGCTGGGAGAACTTTGTA 5058

RESULT 14
LOCUS BD105986 5098 bp DNA linear PAT 18-SEP-2002
DEFINITION Novel LDL-receptor.
ACCESSION BD105986
VERSION BD105986.1 GI:23200804
KEYWORDS JP 2002501376-A/1.
SOURCE Chlamydia sp.
ORGANISM Bacteria; Chlamydiae; Chlamydiaceae; Chlamydia.
REFERENCE 1 (bases 1 to 5098)
AUTHORS Todd, J.A., Hess, J.W., Caskey, C.T., Cox, R.D., Gerhold, D., Hammond, H.
and Hey, P.
TITLE Novel LDL-receptor
JOURNAL Patent: JP 2002501376-A 1 15-JAN-2002;
THE WELLCOME TRUST LTD AS TRUSTEE TO THE WELLCOME TRUST, MERCK & CO
INC
COMMENT PN JP 2002501376-A/1
PD 15-JAN-2002
PF 15-APR-1998 JP 1998543635
PR 15-APR-1997 US 60/043553 05-JUN-1997 US 60/048740 PI
JOHN ANDREW TODD, JOHN WILFRED HESS, CHARLES
THOMAS CASKEY, ROGER
PI DAVID COX,
PC C12N15/12, C12N15/11, C12Q1/68, C07K14/705, C07K16/28, A61K38/17,
PC A61K39/395,
PC A61K48/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source 1. 5098
/organism="Chlamydia sp."
/mol_type="genomic DNA"
/db_xref="taxon:35827"
ORIGIN
Query Match 99.3%; Score 5028; DB 6; Length 5098;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 5053; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
QY 4 ATGAGACCCGAGTAGCGCGCGGGCCCGTCCGGCCGCGGCAACATGAGGCGACGC 63
Db 1 ATGAGACCCGAGTAGCGCGCGGGCCCGTCCGGCCGCGGCAACATGAGGCGACGC 60
QY 64 CCGCCCGGGCCGCGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCGGC 123
Db 61 CCGCCCGGGCCGCGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCGGC 120
QY 124 TGGCGGGCCCGCGCGGCTCGCGCTCTGCTATTTGCAACCGCGGAGCTACGG 183
Db 121 TGGCGGGCCCGCGCGGCTCGCGCTCTGCTATTTGCAACCGCGGAGCTACGG 180
QY 184 CTGGTGACGCGCGGAGTCAAGCTGAGTCCACCATCGTGGTCAAGCGGCTGGAGAT 243
```

```
Db 181 CTGGTGACGCGCGGAGTCAAGCTGAGTCCACCATCGTGGTCAAGCGGCTGGAGAT 240
QY 244 CCGCCCGAGTGGACTTCAGTCTTCCAGGGAGCCGTGTACTGGACAGAGCTGAGCGAG 303
Db 241 CCGCCCGAGTGGACTTCAGTCTTCCAGGGAGCCGTGTACTGGACAGAGCTGAGCGAG 300
QY 304 GAGGCCATCAAGCAGACTCTCTCAACAGAGCGGGCCCGCTGCGAGAACTGGTCAATC 363
Db 301 GAGGCCATCAAGCAGACTCTCTCAACAGAGCGGGCCCGCTGCGAGAACTGGTCAATC 360
QY 364 TCCGGCTGTGTCTCTCCCGAGCGGCTCTGCTGCGACTGGGTGGGCAAGAGCTGTACTGG 423
Db 361 TCCGGCTGTGTCTCTCCCGAGCGGCTCTGCTGCGACTGGGTGGGCAAGAGCTGTACTGG 420
QY 424 ACGGACTCAGAGACAAACCGCATCGAGGTGGCCAACTCAATGGCAATCCCGGAGGTG 483
Db 421 ACGGACTCAGAGACAAACCGCATCGAGGTGGCCAACTCAATGGCAATCCCGGAGGTG 480
QY 484 CTCTTCTGGCAGGACCTTTGACAGCCTAGGGCCATCGCCTTTGGACCCCGCTCACGGGTAC 543
Db 481 CTCTTCTGGCAGGACCTTTGACAGCCTAGGGCCATCGCCTTTGGACCCCGCTCACGGGTAC 540
QY 544 ATGTACTTGGACAGACTGGGGTGGACAGCCCGCGGATTGAGCGGGCAGGGATGGTGCACGC 603
Db 541 ATGTACTTGGACAGACTGGGGTGGACAGCCCGCGGATTGAGCGGGCAGGGATGGTGCACGC 600
QY 604 ACCCGAAGATCATTTGGACTCGGACATTTACTTGGGCCAATGAGTACCATGACCATG 663
Db 601 ACCCGAAGATCATTTGGACTCGGACATTTACTTGGGCCAATGAGTACCATGACCATG 660
QY 664 GAGGAGCAGAGCTCTACTTGGGCTGACGCGCAAGCTCAGCTTCATCCACCGTCCCAACCTG 723
Db 661 GAGGAGCAGAGCTCTACTTGGGCTGACGCGCAAGCTCAGCTTCATCCACCGTCCCAACCTG 720
QY 724 GACGGCTCGTTCCGGCAGAAAGTGGTGGAGGGCAGCTGACGCAACCCCTTCGCCCTGACG 783
Db 721 GACGGCTCGTTCCGGCAGAAAGTGGTGGAGGGCAGCTGACGCAACCCCTTCGCCCTGACG 780
QY 784 CTCTCTCGGGGACACTCTGTACTTGGACAGACTGGCAGACCCCGCTCCATCCATGCGTGAAC 843
Db 781 CTCTCTCGGGGACACTCTGTACTTGGACAGACTGGCAGACCCCGCTCCATCCATGCGTGAAC 840
QY 844 AAGCGCACTGGGGGAGAGAGAGAGAGTCTGTAGTGGCTCTACTTCAACCCATGAGCATC 903
Db 841 AAGCGCACTGGGGGAGAGAGAGAGAGTCTGTAGTGGCTCTACTTCAACCCATGAGCATC 900
QY 904 CAGGTGTGAGCCAGGAGCGGACGCTTTCTTCCACACTCGCTGTGAGGAGGACAAATGGC 963
Db 901 CAGGTGTGAGCCAGGAGCGGACGCTTTCTTCCACACTCGCTGTGAGGAGGACAAATGGC 960
QY 964 GGCTGTCTCCACCTGTGCTGTCTCCCAAGCGAGCTTTCTACACATGCGCTGCCCC 1023
Db 961 GGCTGTCTCCACCTGTGCTGTCTCCCAAGCGAGCTTTCTACACATGCGCTGCCCC 1020
QY 1024 ACGGCTGTGAGCTGAGGACCAACGCGAGAGCTGTAAAGGAGGAGCGGAGAGGTGCTG 1083
Db 1021 ACGGCTGTGAGCTGAGGACCAACGCGAGAGCTGTAAAGGAGGAGCGGAGAGGTGCTG 1080
QY 1084 CTGCTGCCCGGGCGGACCGACTTACGAGGATTTCTGCTGGACACGCGCGGATTTCAACCGAC 1143
Db 1081 CTGCTGCCCGGGCGGACCGACTTACGAGGATTTCTGCTGGACACGCGCGGATTTTACCGAC 1140
QY 1144 ATCGTGTGAGGAGGAGGACATCCGCGACGCGCATTTGCCATCGATAGACCCGCTAGAG 1203
Db 1141 ATCGTGTGAGGAGGAGGACATCCGCGACGCGCATTTGCCATCGATAGACCCGCTAGAG 1200
QY 1204 GGCTATGTCTTACTTGGACAGATGACGAGGTGGGGCCATCCGAGGGCTACTCTGAGCGGG 1263
Db 1201 GGCTATGTCTTACTTGGACAGATGACGAGGTGGGGCCATCCGAGGGGCTACTCTGAGCGGG 1260
QY 1264 TCTGGGGCGGAGCGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGGTGCAC 1323
```

Db 1261 TCTGGGGCGCAGACGCTGTGTCACACCGAGATCAACGACCCCGATGGCATCGGGTCGAC 1320
QY 1324 TGGGTGGCCCGGAAACCTCTACTGGACCGACACGCGGACCGGACCTGAGGTGACGCGC 1383
Db 1321 TGGGTGGCCCGGAAACCTCTACTGGACCGACACGCGGACCGGACCTGAGGTGACGCGC 1380
QY 1384 CTCACGCGCACCTCCCGCAAGATCTCTGTGTGGAGGACCTGAGAGCCCGGAGCCATC 1443
Db 1381 CTCACGCGCACCTCCCGCAAGATCTCTGTGTGGAGGACCTGAGAGCCCGGAGCCATC 1440
QY 1444 GCACCTGCACCCCGTGATGGGCTCTATGTACTGGACAGACTGGGGAGAGAACCTTAAATC 1503
Db 1441 GCACCTGCACCCCGTGATGGGCTCTATGTACTGGACAGACTGGGGAGAGAACCTTAAATC 1500
QY 1504 GAGTGTGCCAACTTGGATGGGACGAGCGGCTGTGTGCTGGTCAATGCTCCCTCGGGTGG 1563
Db 1501 GAGTGTGCCAACTTGGATGGGACGAGCGGCTGTGTGCTGGTCAATGCTCCCTCGGGTGG 1560
QY 1564 CCCAACGGGCTGGCCCTGGACCTGACGAGGAGGAAAGCTCTACTGGGGAGAGCCCAAGACA 1623
Db 1561 CCCAACGGGCTGGCCCTGGACCTGACGAGGAGGAAAGCTCTACTGGGGAGAGCCCAAGACA 1620
QY 1624 GACAGATCGAGGTGATCAATTTGATGGGACGAGAGCGGACCTCTCTGGAGGACAAAG 1683
Db 1621 GACAGATCGAGGTGATCAATTTGATGGGACGAGAGCGGACCTCTCTGGAGGACAAAG 1680
QY 1684 CTCGCGCACATTTTGGGTTCACGCTGTGGGGGACTTCACTACTGACTGACTGGGAG 1743
Db 1681 CTCGCGCACATTTTGGGTTCACGCTGTGGGGGACTTCACTACTGACTGACTGGGAG 1740
QY 1744 CGCCGACGATCGAGCGGGTGCAAGGTCAAGGCCAGCGGAGAGTCACTCATCTTACCCAG 1803
Db 1741 CGCCGACGATCGAGCGGGTGCAAGGTCAAGGCCAGCGGAGAGTCACTCATCTTACCCAG 1800
QY 1804 CTGCGCGACCTGATGGGGCTCAAAGCTGTGAATGTGGCCAGGTCTGGAAACCAACCG 1863
Db 1801 CTGCGCGACCTGATGGGGCTCAAAGCTGTGAATGTGGCCAGGTCTGGAAACCAACCG 1860
QY 1864 TGTGGGACAGAAACGGGGGTGACCCACCTGTCTTCTTCAACCCCGACGACCCGG 1923
Db 1861 TGTGGGACAGAAACGGGGGTGACCCACCTGTCTTCTTCAACCCCGACGACCCGG 1920
QY 1924 TGTGGCTGCCCATCGGCTGGAGCTGTGTGATGACATGAAGACCTGCTGTGCTCTGAG 1983
Db 1921 TGTGGCTGCCCATCGGCTGGAGCTGTGTGATGACATGAAGACCTGCTGTGCTCTGAG 1980
QY 1984 GCCTTCTTGTGTTTCAACAGACAGCGCGCATCCACAGATCTCCCTCGAGACCAATTAAC 2043
Db 1981 GCCTTCTTGTGTTTCAACAGACAGCGCGCATCCACAGATCTCCCTCGAGACCAATTAAC 2040
QY 2044 AACGACGTGGCCATCCGCTCACGGGCGTCAAGGAGGCTCAGCCCTGGAATTTGATGTG 2103
Db 2041 AACGACGT- GCCATCCGCTCACGGGCGTCAAGGAGGCTCAGCCCTGGAATTTGATGTG 2099
QY 2104 TCCAAACACCAATCTACTGGACAGACGCTCAGCCTGGAAGACCATCAGCGCGCTTCATG 2163
Db 2100 TCCAAACACCAATCTACTGGACAGACGCTCAGCCTGGAAGACCATCAGCGCGCTTCATG 2159
QY 2164 AACGGGAGCTCGGTGGAGCACTGTGTGAGTTTGGCTTGAATACCCGAGGGGATGGCC 2223
Db 2160 AACGGGAGCTCGGTGGAGCACTGTGTGAGTTTGGCTTGAATACCCGAGGGGATGGCC 2219
QY 2224 GTTGACTGGATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTG 2283
Db 2220 GTTGACTGGATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTG 2279
QY 2284 GCGGGCTGGACGGGAGTTCCGGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCCGAGG 2343
Db 2280 GCGGGCTGGACGGGAGTTCCGGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCCGAGG 2339
QY 2344 TCGCTGGCCCTGGATCCCAACGAGCTTACATCTACTGGACGAGTGGGGCGGACCGG 2403
Db 2340 TCGCTGGCCCTGGATCCCAACGAGGCTTACATCTACTGGACGAGTGGGGCGGACCGG 2399

QY 2404 AGGATCGTGGGGCCTTTCATGGACCGGACCAACTGCATGACGCTGGTGGACAAAGGTGGGC 2463
Db 2400 AGGATCGTGGGGCCTTTCATGGACCGGACCAACTGCATGACGCTGGTGGACAAAGGTGGGC 2459
QY 2464 CGGGCCAAAGACCTCACCATTTGACTACGCTGACAGCGCTCTACTGGACCGACCTGGAC 2523
Db 2460 CGGGCCAAAGACCTCACCATTTGACTACGCTGACAGCGCTCTACTGGACCGACCTGGAC 2519
QY 2524 ACCAAATGATGATCGAGTCTGCTCAACATGCTGGGTGACGAGCGGCTGCTGATTCGCCACGAT 2583
Db 2520 ACCAAATGATGATCGAGTCTGCTCAACATGCTGGGTGACGAGCGGCTGCTGATTCGCCACGAT 2579
QY 2584 CTCGCGCACCCGTTCTGCTGTGACGAGTACGAGGATTAATCTACTGGACAGACTGGAAT 2643
Db 2580 CTCGCGCACCCGTTCTGCTGTGACGAGTACGAGGATTAATCTACTGGACAGACTGGAAT 2639
QY 2644 CTGACACGATTTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCTCATCCAGGCG 2703
Db 2640 CTGACACGATTTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCTCATCCAGGCG 2699
QY 2704 CACTTGGACTTTCGTGATGGACATCTCTGTGTTCACCTCTCCCGCAGGATGGCTCAAT 2763
Db 2700 CACTTGGACTTTCGTGATGGACATCTCTGTGTTCACCTCTCCCGCAGGATGGCTCAAT 2759
QY 2764 GACTGTATGACAAACAGGGGAGTGTGGGAGCTGTGCTTGCATCCCGCGGCGGCGAC 2823
Db 2760 GACTGTATGACAAACAGGGGAGTGTGGGAGCTGTGCTTGCATCCCGCGGCGGCGAC 2819
QY 2824 CGCTGGGCTCGGCTCACACTACACCTTGACCCAGCAGCGCAACTGACGCGCGGCC 2883
Db 2820 CGCTGGGCTCGGCTCACACTACACCTTGACCCAGCAGCGCAACTGACGCGCGGCC 2879
QY 2884 ACCACCTTTCGTGCTTTCAGCCAGAAATCTCCATTCAGTCCGATGATCCCGAGACCCAG 2943
Db 2880 ACCACCTTTCGTGCTTTCAGCCAGAAATCTCCATTCAGTCCGATGATCCCGAGACCCAG 2939
QY 2944 CACAGCCGGATCTCATCTCGCCCTGCATCGGATGAGGACGTCGAAGCCATCGACTAT 3003
Db 2940 CACAGCCGGATCTCATCTCGCCCTGCATCGGATGAGGACGTCGAAGCCATCGACTAT 2999
QY 3004 GACCACTTGGACAAAGTTTCATCTACTGGGTGATGGGCGGACCAACATCAAGCGAGCCAG 3063
Db 3000 GACCACTTGGACAAAGTTTCATCTACTGGGTGATGGGCGGACCAACATCAAGCGAGCCAG 3059
QY 3064 GACGACGGGACCCAGCCCTTTTGTGACCTCTCTGAGCCAAAGGCCAAAAACCGAGACGG 3123
Db 3060 GACGACGGGACCCAGCCCTTTTGTGACCTCTCTGAGCCAAAGGCCAAAAACCGAGACGG 3119
QY 3124 CAGCCCGACGACCTCAGCATCGACATCTACAGCGGACACCTGTTCTGGACGTGGAGGCC 3183
Db 3120 CAGCCCGACGACCTCAGCATCGACATCTACAGCGGACACCTGTTCTGGACGTGGAGGCC 3179
QY 3184 ACCAATACCATCAACGCTCAGCGGCGGAAAGCCATGGGGGTGGTCTGCTGGGG 3243
Db 3180 ACCAATACCATCAACGCTCAGCGGCGGAAAGCCATGGGGGTGGTCTGCTGGGG 3239
QY 3244 GACCGGACAAAGCCAGGGCCATCTGTGTGAAACCGGAGCGAGGTTACTGTACTTACCC 3303
Db 3240 GACCGGACAAAGCCAGGGCCATCTGTGTGAAACCGGAGCGAGGTTACTGTACTTACCC 3299
QY 3304 AACATGAGGACCGGGCAGCCAAAGATCGAACGCGACCTTGGACCGGACCGAGCGCGAG 3363
Db 3300 AACATGAGGACCGGGCAGCCAAAGATCGAACGCGACCTTGGACCGGACCGAGCGCGAG 3359
QY 3364 GTCTCTTTCACCAACCGGCTCATCCGCTCTGTGGCCCTGTGGCTGTGGTGGACAAACACTGGGC 3423
Db 3360 GTCTCTTTCACCAACCGGCTCATCCGCTCTGTGGCCCTGTGGCTGTGGTGGACAAACACTGGGC 3419
QY 3424 AAGCTGTCTGGGTGGAGCGGAGCTGAAGCGGATTTGAGACTGTGACTGTTCAGGGGCC 3483
Db 3420 AAGCTGTCTGGGTGGAGCGGAGCTGAAGCGGATTTGAGACTGTGACTGTTCAGGGGCC 3479

Db 121 TGCCGGCCCCCGCGCGCTCGCGCTCTCTGCTATTGTTGCCAACCGCGGGACGTACGG 180
Qy 184 CTGTTGGACGCGCGGCGGAGTCAAGCTGAGTCCACCATCGTGGTTCAGCGGCTGGAGAT 243
Db 181 CTGTTGGACGCGCGGCGGAGTCAAGCTGAGTCCACCATCGTGGTTCAGCGGCTGGAGAT 240
Qy 244 GCGCGCGAGTGGACTTCAGTTTTCAGAGGAGCGGTGTTACTGGACAGAGTGGAGCGAG 303
Db 241 GCGCGCGAGTGGACTTCAGTTTTCAGAGGAGCGGTGTTACTGGACAGAGTGGAGCGAG 300
Qy 304 GAGGCACTCAACAGACACTCACTGAAACAGACGCGGCGCGCTGTCAGAACGTGTCTATC 363
Db 301 GAGGCACTCAACAGACACTCACTGAAACAGACGCGGCGCGCTGTCAGAACGTGTCTATC 360
Qy 364 TCCGCGCTGGTCTCTCCGACGCGCTCGCTGCGACTGGGTGGGCAAGAGTGTACTGG 423
Db 361 TCCGCGCTGGTCTCTCCGACGCGCTCGCTGCGACTGGGTGGGCAAGAGTGTACTGG 420
Qy 424 ACGGACTCAGACCAACCGCATCGAGTGGCCAACTCAATGGCACATCCCGGAAGGTG 483
Db 421 ACGGACTCAGACCAACCGCATCGAGTGGCCAACTCAATGGCACATCCCGGAAGGTG 480
Qy 484 CTCCTTCTGCGAGACTTGTACAGCTTAGGGCCATCGCTTGGACCCCGCTCA CGGGTAC 543
Db 481 CTCCTTCTGCGAGACTTGTACAGCTTAGGGCCATCGCTTGGACCCCGCTCA CGGGTAC 540
Qy 544 ATGTACTGACAGACTGGGGTGAGACGCGCCCGGATTTAGCGGGCAGGATGGACG 603
Db 541 ATGTACTGACAGACTGGGGTGAGACGCGCCCGGATTTAGCGGGCAGGATGGATGGACG 600
Qy 604 ACCCGGAGATCAATTGTGACCTCGACATTTACTGSCCAATGGACTGACCATCGACCTG 663
Db 601 ACCCGGAGATCAATTGTGACCTCGACATTTACTGSCCAATGGACTGACCATCGACCTG 660
Qy 664 GAGGAGCAGAGCTTACTTGGGCTGACGCCAAGCTCAGCTTCACTCCACGCTGCCAACCTG 723
Db 661 GAGGAGCAGAGCTTACTTGGGCTGACGCCAAGCTCAGCTTCACTCCACGCTGCCAACCTG 720
Qy 724 GACGGCTCTTCCGGCAGAGTGTGGAGGGCAGCTGACGCCACCCCTTCCGCTGAG 783
Db 721 GACGGCTCTTCCGGCAGAGTGTGGAGGGCAGCTGACGCCACCCCTTCCGCTGAG 780
Qy 784 CTCCTCGGAGACACTCTGTACTGGACAGACTGGCAGCCGCTCCATCATGCTGCAAC 843
Db 781 CTCCTCGGAGACACTCTGTACTGGACAGACTGGCAGCCGCTCCATCATGCTGCAAC 840
Qy 844 AAGCGCACTGGGGGGAAGAGATCTGTAGTGGCTTACTCAACCCATGGACATC 903
Db 841 AAGCGCACTGGGGGGAAGAGATCTGTAGTGGCTTACTCAACCCATGGACATC 900
Qy 904 CAGGTGCTGAGCCAGGAGCGGACCTTTCTTCCACACTCGCTGTGAGGAGCAATGGC 963
Db 901 CAGGTGCTGAGCCAGGAGCGGACCTTTCTTCCACACTCGCTGTGAGGAGCAATGGC 960
Qy 964 GGTGCTGCCACTGTGCTGTGCTCCCAAGCGAGCCTTTTACACATGCGCTGCCCC 1023
Db 961 GGTGCTGCCACTGTGCTGTGCTCCCAAGCGAGCCTTTTACACATGCGCTGCCCC 1020
Qy 1024 ACGGGTGTGAGCTGACGACAAACGACAGAGCTGTAAAGGAGGAGCCGAGAGGTGCTG 1083
Db 1021 ACGGGTGTGAGCTGACGACAAACGACAGAGCTGTAAAGGAGGAGCCGAGAGGTGCTG 1080
Qy 1084 CTGCTGGCGCGGAGCGGACCTTACGGAGATCTCGCTGGACAGCGCGGACTTCAACGAC 1143
Db 1081 CTGCTGGCGCGGAGCGGACCTTACGGAGATCTCGCTGGACAGCGCGGACTTCAACGAC 1140
Qy 1144 ATCGTGTGAGTGGACGACATCCGCGACCGCATTTGCCATCGACTACGACCGCTAGAG 1203
Db 1141 ATCGTGTGAGTGGACGACATCCGCGACCGCATTTGCCATCGACTACGACCGCTAGAG 1200
Qy 1204 GGTATGTCTTACTGGACAGATGAGAGTGGGGCCATCCGAGGGGCTACTTGGACGGG 1263
Db 1201 GGTATGTCTTACTGGACAGATGAGAGTGGGGCCATCCGAGGGGCTACTTGGACGGG 1260

Qy 1264 TCTGGGCGCAGAGCTGTGTTCAACACCGAGATCAACGACCCCGATGTCATCGCGGTTCGAC 1323
Db 1261 TCTGGGCGCAGAGCTGTGTTCAACACCGAGATCAACGACCCCGATGTCATCGCGGTTCGAC 1320
Qy 1324 TGGGTGCGCCGAAACCTTCTACTGACGACACGCGGACCGGACCTGAGAGTTCGACGCG 1383
Db 1321 TGGGTGCGCCGAAACCTTCTACTGACGACACGCGGACCGGACCTGAGAGTTCGACGCG 1380
Qy 1384 CTCAACGCGACCTTCCGCAAGATCTCTGTTGTCGAGAGACTTGAACGAGCCCGAGCCATC 1443
Db 1381 CTCAACGCGACCTTCCGCAAGATCTCTGTTGTCGAGAGACTTGAACGAGCCCGAGCCATC 1440
Qy 1444 GCACTGACACCCCGTGAATGSGGCTCATCTACTGACAGACTGGGAGAGAAACCTTAAATC 1503
Db 1441 GCACTGACACCCCGTGAATGSGGCTCATCTACTGACAGACTGGGAGAGAAACCTTAAATC 1500
Qy 1504 GAGTGTGCCAACTTTGGAATGGGACGAGCGGCTGTGTGTTCAATGCTTCCCTCGGGTGG 1563
Db 1501 GAGTGTGCCAACTTTGGAATGGGACGAGCGGCTGTGTGTTCAATGCTTCCCTCGGGTGG 1560
Qy 1564 CCCAAGCGCTTGGCCCTGGACCTGACAGAGGGGAGCTTACTTGGGAGAGACCCCAAGACA 1623
Db 1561 CCCAAGCGCTTGGCCCTGGACCTGACAGAGGGGAGCTTACTTGGGAGAGACCCCAAGACA 1620
Qy 1624 GACAAGATCGAGGTGATCAATGTTGATGGGACAGAGGGCGGACCTCTCTGGAGGACAAG 1683
Db 1621 GACAAGATCGAGGTGATCAATGTTGATGGGACAGAGGGCGGACCTCTCTGGAGGACAAG 1680
Qy 1684 CTCGCGCACATTTTTCGGGTTCACTGCTGGGGACCTTCTACTTGAAGTGTGAGCTGGCAG 1743
Db 1681 CTCGCGCACATTTTTCGGGTTCACTGCTGGGGACCTTCTACTTGAAGTGTGAGCTGGCAG 1740
Qy 1744 CGCGCAGCATCGAGCGGCTGCAAGGTCAAGGTCGAGCGGAGCTCATCATTTGACCGAG 1803
Db 1741 CGCGCAGCATCGAGCGGCTGCAAGGTCAAGGTCGAGCGGAGCTCATCATTTGACCGAG 1800
Qy 1804 CTGCGCGACCTGATGGGGCTCAAAAGCTGTGAATGTGGCCAAAGTCTGTCGGAACCAACCG 1863
Db 1801 CTGCGCGACCTGATGGGGCTCAAAAGCTGTGAATGTGGCCAAAGTCTGTCGGAACCAACCG 1860
Qy 1864 TGTGCGGACAGAAACGCGGGGTGAGCCACTGTGCTTCTTTCACACCCCAACGACCCGG 1923
Db 1861 TGTGCGGACAGAAACGCGGGGTGAGCCACTGTGCTTCTTTCACACCCCAACGACCCGG 1920
Qy 1924 TGTGCTGCCCATCGGCTGAGCTGAGTGCATGAAGAAGCTGCTGCTGAG 1983
Db 1921 TGTGCTGCCCATCGGCTGAGCTGAGTGCATGAAGAAGCTGCTGCTGAG 1980
Qy 1984 GCCTTCTTGGTCTTTCACGACAGAGCCGCTTCAACAGGATCTCCCTCGAGACCAATTAAC 2043
Db 1981 GCCTTCTTGGTCTTTCACGACAGAGCCGCTTCAACAGGATCTCCCTCGAGACCAATTAAC 2040
Qy 2044 AACGAGTGGCCATCCCGCTCAAGGGCTCAAGGGCTCAGCCCTGAGCTTTGATCTG 2103
Db 2041 AACGAGTGGCCATCCCGCTCAAGGGCTCAAGGGCTCAGCCCTGAGCTTTGATCTG 2100
Qy 2104 TCCAAACACCATCTACTGAGACAGCTGAGCTTGAAGACCATCAGCGCGCTTCTGATG 2163
Db 2101 TCCAAACACCATCTACTGAGACAGCTGAGCTTGAAGACCATCAGCGCGCTTCTGATG 2160
Qy 2164 AACCGGAGCTCGTGGAGCAGTGGTGGAGTTTGGCTTGAATACCCGAGGGCATGGCC 2223
Db 2161 AACCGGAGCTCGTGGAGCAGTGGTGGAGTTTGGCTTGAATACCCGAGGGCATGGCC 2220
Qy 2224 GTTGA CTGGA TGGGCAAGAACTTCTACTTGGGCGGACCTGGGACCAACAGAAATCCGAAGT 2283
Db 2221 GTTGA CTGGA TGGGCAAGAACTTCTACTTGGGCGGACCTGGGACCAACAGAAATCCGAAGT 2280
Qy 2284 GCGCGGCTGGACGGGCAAGTTCGCGCAAGTCTCTGCTGTGGAGGGGACTTGAACACCGGAG 2343
Db 2281 GCGCGGCTGGACGGGCAAGTTCGCGCAAGTCTCTGCTGTGGAGGGGACTTGGACCAACCGGAG 2340

[illegible]

Db	3421	CGCCCTGTGGCCCTTG3TGGTGGACAAACACACTGGGCAAGCTGTTCTGGGTGGAGCGGAC	3480
Qy	3448	CTGAAGCGCATTTGAGAGCTGTGACCTCTGTGAGGGCCAAACCGCTGACCCCTGGAGGAGCGCC	3507
Db	3481	CTGAAGCGCATTTGAGAGCTGTGACCTCTGTGAGGGCCAAACCGCTGACCCCTGGAGGAGCGCC	3540
Qy	3508	AACATCGTGCAGCCTCTTGGGCTTGACCATCTTTGGCAAGCATCTCTACTGGATCGACCGC	3567
Db	3541	AACATCGTGCAGCCTCTTGGGCTTGACCATCTTTGGCAAGCATCTCTACTGGATCGACCGC	3600
Qy	3568	CAGCAGCAGATGATCGAGCGTGTGGAGAAAGACACACCGGGGACAAGCGGACTTCGATCCAG	3627
Db	3601	CAGCAGCAGATGATCGAGCGTGTGGAGAAAGACACACCGGGGACAAGCGGACTTCGATCCAG	3660
Qy	3628	GGCCGTGTGGCCACCTCACTTGGCATCCATGTGAGTGGAGGAAGTCAAGCTTGGAGGAGTTC	3687
Db	3661	GGCCGTGTGGCCACCTCACTTGGCATCCATGTGAGTGGAGGAAGTCAAGCTTGGAGGAGTTC	3720
Qy	3688	TCAGCCCAACCATGTGCCCGGTGCAAAATGSGGTGCTCCACATCTGTATTGCCAAGGGT	3747
Db	3721	TCAGCCCAACCATGTGCCCGGTGCAAAATGSGGTGCTCCACATCTGTATTGCCAAGGGT	3780
Qy	3748	GATGGGACACACCGGTGCTCATGCCAGTCCACCTCGTGTCTCTGCAGAACCTTGTCTGACC	3807
Db	3781	GATGGGACACACCGGTGCTCATGCCAGTCCACCTCGTGTCTCTGCAGAACCTTGTCTGACC	3840
Qy	3808	TGTGGAGAGCGCCACCTGCTCCCGGACAGTTTGTGCATGTGCCACAGGGGAGATCGAC	3867
Db	3841	TGTGGAGAGCGCCACCTGCTCCCGGACAGTTTGTGCATGTGCCACAGGGGAGATCGAC	3900
Qy	3868	TGTATCCCCCGGGGCTTGGCGCTGTGACGGCTTTTCCGAGTGCATGACACAGAGCGACGAG	3927
Db	3901	TGTATCCCCCGGGGCTTGGCGCTGTGACGGCTTTTCCGAGTGCATGACACAGAGCGACGAG	3960
Qy	3928	GAGGGCTGCCCGTGTGCTCTCGCGGCCAGTTTCCCTTGCAGGCGGGGTCAAGTGTGTGGAC	3987
Db	3961	GAGGGCTGCCCGTGTGCTCTCGCGGCCAGTTTCCCTTGCAGGCGGGGTCAAGTGTGTGGAC	4020
Qy	3988	CTGCGCTGCGCTGCGACGCGGAGGCGAGCTGTTCAGGACCGCTTCAGACGAGGCGGACTGT	4047
Db	4021	CTGCGCTGCGCTGCGACGCGGAGGCGAGCTGTTCAGGACCGCTTCAGACGAGGCGGACTGT	4080
Qy	4048	GACGCCATCTGCTTCCCAACCAAGTTTCCGCTGTGCGAGCGGCCAGTGTCTCATCAAA	4107
Db	4081	GACGCCATCTGCTTCCCAACCAAGTTTCCGCTGTGCGAGCGGCCAGTGTCTCATCAAA	4140
Qy	4108	CAGCAGTGGCATCTCTTTCGCCGACTGTATCGAGGCTTCGACGAGCTCATGTGTGAATC	4167
Db	4141	CAGCAGTGGCATCTCTTTCGCCGACTGTATCGAGGCTTCGACGAGCTCATGTGTGAATC	4200
Qy	4168	ACCAAGCGGCCCTCAGACGACAGCCCGGCCACAGCAGATGCCATCGGGCCCGTCATTGGC	4227
Db	4201	ACCAAGCGGCCCTCAGACGACAGCCCGGCCACAGCAGATGCCATCGGGCCCGTCATTGGC	4260
Qy	4228	ATCATCTCTCTCTTCGCTCATGGGTGTGTCTATTTTGTGTGCCAGCGCTGTGTGTGC	4287
Db	4261	ATCATCTCTCTCTTCGCTCATGGGTGTGTCTATTTTGTGTGCCAGCGCTGTGTGTGC	4320
Qy	4288	CAGCGCTATCGGGGGCAAACGGGCCCTTCCCGACGAGTATGTACGCGGGAACCCCGCAC	4347
Db	4321	CAGCGCTATCGGGGGCAAACGGGCCCTTCCCGACGAGTATGTACGCGGGAACCCCGCAC	4380
Qy	4348	GTGCCCCCTCAATTTTCATAGCCCCGGGGGTTTCCACGATGGCCCCCTTCACAGGCATCGCA	4407
Db	4381	GTGCCCCCTCAATTTTCATAGCCCCGGGGGTTTCCACGATGGCCCCCTTCACAGGCATCGCA	4440
Qy	4408	TGCGGAAAGTCCATGATGAGCTCCGTGAGCTCATGTGGGGGGCCGGGGCGGGGTGCCCTTG	4467
Db	4441	TGCGGAAAGTCCATGATGAGCTCCGTGAGCTCATGTGGGGGGCCGGGGCGGGGTGCCCTTG	4500
Qy	4468	TAGACCGGGAACCAAGTTCACAGGGGGCTCGTCCAGCAGCTCGTCCAGCAGCAAGAGCCACG	4527

Db 4501 TACGACCGGAACCAAGTCCAGGGGCTCGTCCAGCAGCTCGTCCAGCACGAAGCCACG 4560
QY 4528 CTGTACCCGCGATCCTGAAACCGCGCGGCTCCCGGCCACGGACCCCTCCCTGTACAAC 4587
Db 4561 CTGTACCCGCGATCCTGAAACCGCGCGGCTCCCGGCCACGGACCCCTCCCTGTACAAC 4620
QY 4588 ATGGACATGTTCTACTCTTTCAAAACATTCCGGCCACTCGGAGACCGGTACAGGCCCTTACATC 4647
Db 4621 ATGGACATGTTCTACTCTTTCAAAACATTCCGGCCACTCGGAGACCGGTACAGGCCCTTACATC 4680
QY 4648 ATTCGAGGAATGGCGCCCGCCCGACGACGCGCTTCGACACCGAGTGTGTGACAGCGACTAC 4707
Db 4681 ATTCGAGGAATGGCGCCCGCCCGACGACGCGCTTCGACACCGAGTGTGTGACAGCGACTAC 4740
QY 4708 AGCGCCAGCGCTGGAAGCCCGACGAGTACTACCTGGATTGAACTCGGACTCAGACCCC 4767
Db 4741 AGCGCCAGCGCTGGAAGCCCGACGAGTACTACCTGGATTGAACTCGGACTCAGACCCC 4800
QY 4768 TATCCACCCCGCACCCACCGCCACAGCCAGTACTGTCCGGGAGGACAGCTGCCCGGCC 4827
Db 4801 TATCCACCCCGCACCCACCGCCACAGCCAGTACTGTCCGGGAGGACAGCTGCCCGGCC 4860
QY 4828 TCGCCCGCCACCGAGAGAGTACTTCCATCTTTCCCGCCCGCTCCGTCCCGCTGCACG 4887
Db 4861 TCGCCCGCCACCGAGAGAGTACTTCCATCTTTCCCGCCCGCTCCGTCCCGCTGCACG 4920
QY 4888 GACTCATCTGACCTCGCGCGGCGCACCTCGGCTTCTGTGCGCCCTGTAAATAGTTTGA 4947
Db 4921 GACTCATCTGACCTCGCGCGGCGCACCTCGGCTTCTGTGCGCCCTGTAAATAGTTTGA 4980
QY 4948 AATATGAACAAAGAAAAATATATTTTATGATTTTAAAAAATAAATAATTTGGGATTTT 5007
Db 4981 AATATGAACAAAGAAAAATATATTTTATGATTTTAAAAAATAAATAATTTGGGATTTT 5040
QY 5008 AAAAATGAGAAATGTGAACCTGTGATGGGTGGGCGAGGCTGGGAGAACTTTGTA 5063
Db 5041 AAAAATGAGAAATGTGAACCTGTGATGGGTGGGCGAGGCTGGGAGAACTTTGTA 5096

Search completed: February 18, 2005, 20:08:38
Job time : 21195 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2005, 01:25:15 ; Search time 49 seconds
(without alignments)

2460.372 Million cell updates/sec

Title: US-09-931-375A-2

Perfect score: 8740

Sequence: 1 MEAAPGPPWPLLLLLLLLL.....TERSYPHLPFPSPCTDSS 1615

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8736	100.0	1615	4	US-09-060-299-3
2	8736	100.0	1615	4	US-09-402-923A-3
3	8715	99.7	1615	4	US-09-544-398B-3
4	8715	99.7	1615	4	US-09-543-771B-3
5	8706	99.6	1615	4	US-09-544-398B-4
6	8706	99.6	1615	4	US-09-543-771B-4
7	8607	98.5	1591	4	US-09-060-299-4
8	8607	98.5	1591	4	US-09-402-923A-4
9	8607	98.5	1591	4	US-09-402-923A-43
10	8607	98.5	1591	4	US-09-060-299-29
11	8574	98.1	1639	4	US-09-402-923A-29
12	8574	98.1	1639	4	US-09-060-299-39
13	8568	98.0	1584	4	US-09-402-923A-39
14	8568	98.0	1584	4	US-09-060-299-42
15	8259.5	94.5	1614	4	US-09-402-923A-42
16	8259.5	94.5	1614	4	US-09-060-299-44
17	8208	93.9	1586	4	US-09-402-923A-44
18	8208	93.9	1586	4	US-09-060-299-25
19	7882	90.2	1451	4	US-09-402-923A-25
20	7882	90.2	1451	4	US-09-402-923A-25
21	6153.5	70.4	1621	4	US-09-949-016-8450
22	5442	62.3	1148	4	US-09-949-016-8328
23	2937	33.6	550	4	US-09-060-299-7
24	2937	33.6	550	4	US-09-402-923A-7
25	2645.5	30.3	533	4	US-09-060-299-8
26	2645.5	30.3	533	4	US-09-402-923A-8
27	2193	25.1	4544	1	US-08-469-486-52

Sequence 52, Appl
Sequence 84, Appl
Sequence 84, Appl
Sequence 88, Appl
Sequence 90, Appl
Sequence 86, Appl
Sequence 8985, Ap
Sequence 3, Appl
Sequence 7, Appl
Sequence 40, Appl
Sequence 8313, Ap
Sequence 41833, A
Sequence 2, Appl
Sequence 2, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 9528, Ap
Sequence 4, Appl

28 2193 25.1 4544 2 US-08-469-486-52
29 1876.5 21.5 4654 3 US-08-476-515A-84
30 1876.5 21.5 4655 3 US-08-652-877-84
31 1876.5 21.5 4655 3 US-08-652-877-88
32 1876.5 21.5 4655 3 US-08-652-877-90
33 1872.5 21.4 4655 3 US-08-652-877-86
34 1725.5 19.7 2362 4 US-09-949-016-8985
35 786.5 9.0 2213 1 US-08-727-034-3
36 774 8.9 2214 1 US-08-727-034-7
37 774 8.9 2214 4 US-09-919-039-40
38 694 7.9 1345 4 US-09-949-016-8313
39 688.5 7.9 396 4 US-09-270-767-41833
40 650 7.4 873 1 US-08-393-734-2
41 650 7.4 873 3 US-08-894-489-2
42 648 7.4 846 1 US-08-149-103-3
43 648 7.4 846 1 US-08-451-883-3
44 648 7.4 904 4 US-09-949-016-9528
45 646 7.4 846 1 US-08-149-103-4

ALIGNMENTS

RESULT 1

US-09-060-299-3
; Sequence 3, Application US/09060299
; Patent No. 6545137
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; APPLICANT: Hess, John W
; APPLICANT: Caskey, Charles T
; APPLICANT: Cox, Roger D
; APPLICANT: Gerhold, David
; APPLICANT: Hammond, Holly
; APPLICANT: Hey, Patricia
; APPLICANT: Kawaguchi, Yoshihiko
; APPLICANT: Merriman, Tony R
; APPLICANT: Metzker, Michael L
; TITLE OF INVENTION: No. 6545137el Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6545137th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,299
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/043,553
; FILING DATE: 15-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,740
; FILING DATE: 05-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: B.J.Sadoff
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 620-35
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4091
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1615 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

US-09-060-299-3

Query Match 100.0%; Score 8736; DB 4; Length 1615;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MEAPPGPPWPLLLLLLLLLLALCCGPAAPAAASPLLLFANRRDRLVADAGGKLESTIVVS	60
Db	1	MEAPPGPPWPLLLLLLLLLLALCCGPAAPAAASPLLLFANRRDRLVADAGGKLESTIVVS	60
Qy	61	GLEDAAAVDFGSKGAVVTDVSEAIKOTYLNQTAAGVQNVWISGLVSPDGLACDWGK	120
Db	61	GLEDAAAVDFGSKGAVVTDVSEAIKOTYLNQTAAGVQNVWISGLVSPDGLACDWGK	120
Qy	121	KLYWTDSTNRIEVANLNGTSRKVLFWQDLQDRAIALDPAHGYMYWTDWGETPRIERAG	180
Db	121	KLYWTDSTNRIEVANLNGTSRKVLFWQDLQDRAIALDPAHGYMYWTDWGETPRIERAG	180
Qy	181	MDGSTRKLIIVDSIYWPNGLTIDLEEOKLYWADAKLSFIHRANLDGSPKQKVVEGSLTHP	240
Db	181	MDGSTRKLIIVDSIYWPNGLTIDLEEOKLYWADAKLSFIHRANLDGSPKQKVVEGSLTHP	240
Qy	241	FALTLGSDTLWTDQWTRSIHACNKRGTGKKEILSALYSPMDIQVLSOERQPPFHTRCE	300
Db	241	FALTLGSDTLWTDQWTRSIHACNKRGTGKKEILSALYSPMDIQVLSOERQPPFHTRCE	300
Qy	301	EDNGGCSHLCLLSPSEPPYTACPTGVOLQDNGRTCKAGAEVLLLARRTDLRISLDT	360
Db	301	EDNGGCSHLCLLSPSEPPYTACPTGVOLQDNGRTCKAGAEVLLLARRTDLRISLDT	360
Qy	361	DFTDVLQVDDIRHAIADYDPLGYYVWTDDEVRATRRAYLVLDGSGAQTIVNTREINDPDG	420
Db	361	DFTDVLQVDDIRHAIADYDPLGYYVWTDDEVRATRRAYLVLDGSGAQTIVNTREINDPDG	420
Qy	421	IADVARNLYWTDGTDRIEVTRNGTSRKILSVSEDLDEPRAIALHPVGLMTWTDWGE	480
Db	421	IADVARNLYWTDGTDRIEVTRNGTSRKILSVSEDLDEPRAIALHPVGLMTWTDWGE	480
Qy	481	NPKTECANLDQERRVLVNASLWPNGLALDQEGKLYWGDAKTDKIEVINVDGTGKRTL	540
Db	481	NPKTECANLDQERRVLVNASLWPNGLALDQEGKLYWGDAKTDKIEVINVDGTGKRTL	540
Qy	541	LEDKLPHFPGTLLGDFTYWTDQWRRSIEVHVKKASRDVLIIDOLPDLMLGKAVNVAKV	600
Db	541	LEDKLPHFPGTLLGDFTYWTDQWRRSIEVHVKKASRDVLIIDOLPDLMLGKAVNVAKV	600
Qy	601	GTNFCADRNGCCSHLCFFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRISL	660
Db	601	GTNFCADRNGCCSHLCFFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRISL	660
Qy	661	ETNNNDVAIPLTGVEKASALDFVSNHHIYWTDVSLKTIISAPFNGSSVEHVVEFGLDYP	720
Db	661	ETNNNDVAIPLTGVEKASALDFVSNHHIYWTDVSLKTIISAPFNGSSVEHVVEFGLDYP	720
Qy	721	EGMAVDWNGKLYWADGTGNRIEVARLDGQFQVLRVWRDLNPRSLALDPTKGYIYWTEW	780
Db	721	EGMAVDWNGKLYWADGTGNRIEVARLDGQFQVLRVWRDLNPRSLALDPTKGYIYWTEW	780
Qy	781	GGKPRIIVRAFMDGNTCMTLVKVGRANDLTIDYADQRLYWTDLTNMISSNMLGQERVV	840
Db	781	GGKPRIIVRAFMDGNTCMTLVKVGRANDLTIDYADQRLYWTDLTNMISSNMLGQERVV	840
Qy	841	IADDLPHFPGLTQYSDIYTWDMNLHSIERADKTSGRNRTLIQGHLDVWMDILVPHSSRQ	900
Db	841	IADDLPHFPGLTQYSDIYTWDMNLHSIERADKTSGRNRTLIQGHLDVWMDILVPHSSRQ	900
Qy	901	DGLNDCHNNQCQQLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLLSQKSAISRMI	960
Db	901	DGLNDCHNNQCQQLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLLSQKSAISRMI	960
Qy	961	PDDQHSPLILPLHGLRNVAIDYDPLDKFYIYWDGQONIKRAKDDGTQPPVLTSLSQGG	1020
Db	961	PDDQHSPLILPLHGLRNVAIDYDPLDKFYIYWDGQONIKRAKDDGTQPPVLTSLSQGG	1020

Qy	1021	NPDRQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG	1080
Db	1021	NPDRQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG	1080
Qy	1081	LYFTNMQDRAAKIERAALDGTREVLFTTGLIRVALVVDNTLGLFWVDADLKRIBSCD	1140
Db	1081	LYFTNMQDRAAKIERAALDGTREVLFTTGLIRVALVVDNTLGLFWVDADLKRIBSCD	1140
Qy	1141	LSGANRLTLEDANIVQPLGLTILGKLYWIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG	1200
Db	1141	LSGANRLTLEDANIVQPLGLTILGKLYWIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG	1200
Qy	1201	IHAVEEVSLEBFSNHPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQNLITCGEPPTCS	1260
Db	1201	IHAVEEVSLEBFSNHPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQNLITCGEPPTCS	1260
Qy	1261	PDQFACATGETDCIPGAWRCDGPEECDDQSDDEECPCVCSAAQFFCARGQCCVDLRLRCDGE	1320
Db	1261	PDQFACATGETDCIPGAWRCDGPEECDDQSDDEECPCVCSAAQFFCARGQCCVDLRLRCDGE	1320
Qy	1321	ADCCDRSDEADCAICLPNQFRCASGQCVLIKQCCDSFPDCIDGSDDELMECEITKPPSDDS	1380
Db	1321	ADCCDRSDEADCAICLPNQFRCASGQCVLIKQCCDSFPDCIDGSDDELMECEITKPPSDDS	1380
Qy	1381	PAHSSAIGPVIGIITLSLFWMGVYFVQVVCQRYAGANGPPHYYVSGTSPHVLNFIAP	1440
Db	1381	PAHSSAIGPVIGIITLSLFWMGVYFVQVVCQRYAGANGPPHYYVSGTSPHVLNFIAP	1440
Qy	1441	GGSHQGFPTGIACGKSNMSSVSLMGGRGVPLVDRNHTVGTASSSSSSSTKATLYPPLNP	1500
Db	1441	GGSHQGFPTGIACGKSNMSSVSLMGGRGVPLVDRNHTVGTASSSSSSSTKATLYPPLNP	1500
Qy	1501	PPSPATDPSLYNMDMYSSNIPATPYRYIIRGMAPPPTTPTCSTDVCDSDYSASRWKAS	1560
Db	1501	PPSPATDPSLYNMDMYSSNIPATPYRYIIRGMAPPPTTPTCSTDVCDSDYSASRWKAS	1560
Qy	1561	KYLDLNSDSDPYPPPTPHSQVLSAEDSCPPSPATERSYHLEFPSPSPCTDSS	1615
Db	1561	KYLDLNSDSDPYPPPTPHSQVLSAEDSCPPSPATERSYHLEFPSPSPCTDSS	1615

RESULT 2

US-09-402-923A-3
; Sequence 3, Application US/09402923A
; Patent No. 655654

GENERAL INFORMATION:
APPLICANT: Todd, John A

Hess, John W
Caskey, Charles T
Cox, Roger D
Gerhold, David
Hammond, Holly
Hey, Patricia
Kawaguchi, Yoshihiko
Merriman, Tony R
Metzker, Michael L

TITLE OF INVENTION: No. 655654e1 LDL-Receptor
NUMBER OF SEQUENCES: 455
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon and Vanderhye
STREET: 1100 No. 655654th Glebe Road, Eighth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: US
ZIP: VA 22201-4714

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,923A

FILING DATE: 14-Feb-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB98/01102
FILING DATE: 15-APR-1998
APPLICATION NUMBER: US 60/043,553
FILING DATE: 15-APR-1997
APPLICATION NUMBER: US 60/048,740
FILING DATE: 05-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: B.J.Sadoff
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 620-81
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4091
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1615 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-402-923A-3

Query Match 100.0%; Score 8736; DB 4; Length 1615;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEAAPPGPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGGKLESTIVVS 60
DB 1 MEAAPPGPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGGKLESTIVVS 60
QY 61 GLEDAAAVDFGSGAVTWTDSBEAIKQTYLNTGTGAQVNVVSLVSPDGLACDWGK 120
DB 61 GLEDAAAVDFGSGAVTWTDSBEAIKQTYLNTGTGAQVNVVSLVSPDGLACDWGK 120
QY 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLQPAIALDPAHGYMYWTDWGETPRIERAG 180
DB 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLQPAIALDPAHGYMYWTDWGETPRIERAG 180
QY 181 MDGSTRKLIIVDSIYWPNGLIIDLEQKLYWADAKLSFIHRANLDGSPRQKVVGSLTHP 240
DB 181 MDGSTRKLIIVDSIYWPNGLIIDLEQKLYWADAKLSFIHRANLDGSPRQKVVGSLTHP 240
QY 241 FALTLSGDTLWYTDWQTRSIIHACNKRGTGKKEILSALYSMDIQVLQSOERQPFHTRCE 300
DB 241 FALTLSGDTLWYTDWQTRSIIHACNKRGTGKKEILSALYSMDIQVLQSOERQPFHTRCE 300
QY 301 EDNGGCSHLCLLSPSEPPYTCACPTGVQLQDNRTCKAGABEVLLARRTDLRLISLDT 360
DB 301 EDNGGCSHLCLLSPSEPPYTCACPTGVQLQDNRTCKAGABEVLLARRTDLRLISLDT 360
QY 361 DFTDIVLQDDIRHAIADYDPLEGYVYTWDEVRIRRAYLDGSGAQLVNTINPDG 420
DB 361 DFTDIVLQDDIRHAIADYDPLEGYVYTWDEVRIRRAYLDGSGAQLVNTINPDG 420
QY 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALHVPVGLMYWTDWGE 480
DB 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALHVPVGLMYWTDWGE 480
QY 481 NPKIECANLDQERRVLVNASLGNPGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL 540
DB 481 NPKIECANLDQERRVLVNASLGNPGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL 540
QY 541 LEDKLPHFGLTLDGDFIYTWTDWQRRSIEVHVKKASRDVIIDQLPDLMLGLKAVNAKV 600
DB 541 LEDKLPHFGLTLDGDFIYTWTDWQRRSIEVHVKKASRDVIIDQLPDLMLGLKAVNAKV 600
QY 601 GTNPCADRNCGCSHLCPFTPHATCGCPIGLELSDMKTCIVPEAFVFTSRAIHRISL 660
DB 601 GTNPCADRNCGCSHLCPFTPHATCGCPIGLELSDMKTCIVPEAFVFTSRAIHRISL 660
QY 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWTDVSLKTIISRAFMNGSSVEHVVEGLDYP 720
DB 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWTDVSLKTIISRAFMNGSSVEHVVEGLDYP 720

661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWTDVSLKTIISRAFMNGSSVEHVVEGLDYP 720
721 EGMAVDWMGKULYWADTGTNRRIEVARLDGQPRQLVWRDLNPRSLALDPTKGYIYTWTEW 780
721 EGMAVDWMGKULYWADTGTNRRIEVARLDGQPRQLVWRDLNPRSLALDPTKGYIYTWTEW 780
781 GKGPRIVRAFMWGTNCMTLVKVGANDLTIDYADQRLYWTDLTNMIESNMLGQERVV 840
781 GKGPRIVRAFMWGTNCMTLVKVGANDLTIDYADQRLYWTDLTNMIESNMLGQERVV 840
841 IADLPHPPGLTQYSDIYWTDMNLHSTIERADKTSGRNRTLIQGLDFVMDILVFHSSRQ 900
841 IADLPHPPGLTQYSDIYWTDMNLHSTIERADKTSGRNRTLIQGLDFVMDILVFHSSRQ 900
901 DGLNDCHNNGCCGOLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLFQSKAISRM 960
901 DGLNDCHNNGCCGOLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLFQSKAISRM 960
961 PDDQHSPLILPLHGLRNKVIDYDPLDKFYIYWDGRQNIKRAKDDGTQPFVLTSLSQO 1020
961 PDDQHSPLILPLHGLRNKVIDYDPLDKFYIYWDGRQNIKRAKDDGTQPFVLTSLSQO 1020
1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGRDRKPRAIVVNAERG 1080
1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGRDRKPRAIVVNAERG 1080
1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKFWVDADLKRIESC 1140
1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKFWVDADLKRIESC 1140
1141 LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIEVERKTTGDKRTRIQGRVAHLTG 1200
1141 LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIEVERKTTGDKRTRIQGRVAHLTG 1200
1201 IHAVEVSLSEFSAHPCARDNGGCSHICIAKGDGTPRCPCVHLVLLQNLITCGEPPTCS 1260
1201 IHAVEVSLSEFSAHPCARDNGGCSHICIAKGDGTPRCPCVHLVLLQNLITCGEPPTCS 1260
1261 PDQACATGETDCIPGAWRCDFPECDOSDEECPCVCSAAQFCARGQCVDLRLRCDGE 1320
1261 PDQACATGETDCIPGAWRCDFPECDOSDEECPCVCSAAQFCARGQCVDLRLRCDGE 1320
1321 ADCQDRDEADCDAILPNQFRCASGCVLTKQCDSPDCIDGSDLMCMCITKPPSDDS 1380
1321 ADCQDRDEADCDAILPNQFRCASGCVLTKQCDSPDCIDGSDLMCMCITKPPSDDS 1380
1381 PAHSSAIGPVIGIILSLFVMGVYFVQVVCQVYAGANGFPPEHYVSGTPHVPLNFIAP 1440
1381 PAHSSAIGPVIGIILSLFVMGVYFVQVVCQVYAGANGFPPEHYVSGTPHVPLNFIAP 1440
1441 GGSQHGPTGTACGKSNMSSVSLMGGRGVPLVDRNHVTGASSSSSSSTKATLPPILNP 1500
1441 GGSQHGPTGTACGKSNMSSVSLMGGRGVPLVDRNHVTGASSSSSSSTKATLPPILNP 1500
1501 PPSPATPDLNMDMFYSSNIPATVPYPIIRGMAPPPTPCSTDVCDSDYSASRWKAS 1560
1501 PPSPATPDLNMDMFYSSNIPATVPYPIIRGMAPPPTPCSTDVCDSDYSASRWKAS 1560
1561 KYLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLFPFPPSPCTDSS 1615
1561 KYLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLFPFPPSPCTDSS 1615

RESULT 3
US-09-544-398B-3
; Sequence 3, Application US/09544398B
; Patent No. 6770461
; GENERAL INFORMATION:
; APPLICANT: Carulli, John P.
; APPLICANT: Little, Randall D.
; APPLICANT: Recker, Robert R.
; APPLICANT: Johnson, Mark L.
; TITLE OF INVENTION: High bone mass gene of 11q13.3

```
; FILE REFERENCE: 032796-013
; CURRENT APPLICATION NUMBER: US/09/544,398B
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US/09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US/09/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US/09/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 641
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-544-398B-3

Query Match      99.7%; Score 8715; DB 4; Length 1615;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1610; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MEAAPGPPWPLLLLLLLLLLALCGCPAPAAAPLFFANRRDRLVLDAGGVKLESTIVS 60
DB 1 MEAAPGPPWPLLLLLLLLLLALCGCPAPAAAPLFFANRRDRLVLDAGGVKLESTIVS 60
QY 61 GLEDAADVDFQSKGAVYTDVSEBAIKQTYLNOTGAQVNVISGLVSPDGLACDWGK 120
DB 61 GLEDAADVDFQSKGAVYTDVSEBAIKQTYLNOTGAQVNVISGLVSPDGLACDWGK 120
QY 121 KLYWTDSETNRELVANLNGTSKVLFWQDLOPRAIALDPAGHYMYTWDGETPRIERAG 180
DB 121 KLYWTDSETNRELVANLNGTSKVLFWQDLOPRAIALDPAGHYMYTWDGETPRIERAG 180
QY 181 MDGSTRKIIVSDIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPQKVVESGLRHP 240
DB 181 MDGSTRKIIVSDIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPQKVVESGLRHP 240
QY 241 FALTUSGDTLYWTDQWTSIHACNKRGTGKKEIISALYSMDIOVLQERQPPFHTTCE 300
DB 241 FALTUSGDTLYWTDQWTSIHACNKRGTGKKEIISALYSMDIOVLQERQPPFHTTCE 300
QY 301 EDNGGCSHLCLLSPSEPFYTCACPTGVQDNGRTCKAGAEVLLARTRDLRLSLDTP 360
DB 301 EDNGGCSHLCLLSPSEPFYTCACPTGVQDNGRTCKAGAEVLLARTRDLRLSLDTP 360
QY 361 DFTDIVLQVDDIRHAIAIDYDPLEGVYVWTDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
DB 361 DFTDIVLQVDDIRHAIAIDYDPLEGVYVWTDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
QY 421 IAVDVARNLYWTDGTDRIEVTRNGTSKTLVSEDLDEPRALALHPVMGLMYWTDWGE 480
DB 421 IAVDVARNLYWTDGTDRIEVTRNGTSKTLVSEDLDEPRALALHPVMGLMYWTDWGE 480
QY 481 NPKIECANLDGQERRVIVNASLGWPNGLALDIQEGKLYWDAKTDKIEVINVDGTKRRTL 540
DB 481 NPKIECANLDGQERRVIVNASLGWPNGLALDIQEGKLYWDAKTDKIEVINVDGTKRRTL 540
QY 541 LEDKLPHPGFTLLGDFIYVTDQWRRSIRVHKVKAARDVVIDQLPDLMLKAVNVAKV 600
DB 541 LEDKLPHPGFTLLGDFIYVTDQWRRSIRVHKVKAARDVVIDQLPDLMLKAVNVAKV 600
QY 601 GTNPACDRNGGCSHLCCFTPHATRCGCPIGLELLSDMTKCIVPEAPLVFTSRAAHRISL 660
DB 601 GTNPACDRNGGCSHLCCFTPHATRCGCPIGLELLSDMTKCIVPEAPLVFTSRAAHRISL 660
QY 661 ETNNNDVAIPLAGVKEASALDFVSNHHIYTDVSLKTIISRAFNGSSVEHVVEFGLDYP 720
DB 661 ETNNNDVAIPLAGVKEASALDFVSNHHIYTDVSLKTIISRAFNGSSVEHVVEFGLDYP 720
QY 721 EGMADVMMGKNIYADTGTNRILEVARLDGQFQVLRDLNPRSLALDPTKGIYWTW 780
DB 721 EGMADVMMGKNIYADTGTNRILEVARLDGQFQVLRDLNPRSLALDPTKGIYWTW 780
```

```
QY 781 GGPRIIVRAFMGDTNCMTLVKVGANDLTIDYADORLYWTDLTDTNNIESNMLGQERVV 840
DB 781 GGPRIIVRAFMGDTNCMTLVKVGANDLTIDYADORLYWTDLTDTNNIESNMLGQERVV 840
QY 841 IADDLPHFPGLTQVSDIYIYWTDMNLHSIERADKTSGRNRTLIOGHLDLFDVMDILVFHSSRQ 900
DB 841 IADDLPHFPGLTQVSDIYIYWTDMNLHSIERADKTSGRNRTLIOGHLDLFDVMDILVFHSSRQ 900
QY 901 DGLNDCMHNNGCGQLCIAIPGGHRCGCASHYTLDPSSRNCSPPTTLLFSQKSAIRMI 960
DB 901 DGLNDCMHNNGCGQLCIAIPGGHRCGCASHYTLDPSSRNCSPPTTLLFSQKSAIRMI 960
QY 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFYWVDGRQNIKRAKDDGTQFPVLTSLSQGQ 1020
DB 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFYWVDGRQNIKRAKDDGTQFPVLTSLSQGQ 1020
QY 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
DB 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
QY 1081 LYFTNMQDRAAKTBRALDGTREVLFTTGLIRPVALVVDNTLGLKLFWDADLAKRIESED 1140
DB 1081 LYFTNMQDRAAKTBRALDGTREVLFTTGLIRPVALVVDNTLGLKLFWDADLAKRIESED 1140
QY 1141 LSGANRLTLEDAIVQPLGLTILGKHYWIDRQOQMIERVEKTTGDKRTRIQRVAHLTG 1200
DB 1141 LSGANRLTLEDAIVQPLGLTILGKHYWIDRQOQMIERVEKTTGDKRTRIQRVAHLTG 1200
QY 1201 IHAVEVSLSEFSAHPCARDNGGCSHICIAKGDGTGPRCSPVHLVLLQNLITCGEPPTCS 1260
DB 1201 IHAVEVSLSEFSAHPCARDNGGCSHICIAKGDGTGPRCSPVHLVLLQNLITCGEPPTCS 1260
QY 1261 PDQFACATGIDCIPGAWRCDGPECCDDQSDDEECPCVCSAAQFPFCARQCQVDLRLRCDGE 1320
DB 1261 PDQFACATGIDCIPGAWRCDGPECCDDQSDDEECPCVCSAAQFPFCARQCQVDLRLRCDGE 1320
QY 1321 ADCQDRDEADCDCAICLPNCFRCASGQCVLIKQCCDSFPDCIDGSDBLMCEITKPPSDDS 1380
DB 1321 ADCQDRDEADCDCAICLPNCFRCASGQCVLIKQCCDSFPDCIDGSDBLMCEITKPPSDDS 1380
QY 1381 PAHSSAIGPVIGIILSLFVMGGVYVFCQVVCQYAGANGPPHYYSGTTPHPLNFTAP 1440
DB 1381 PAHSSAIGPVIGIILSLFVMGGVYVFCQVVCQYAGANGPPHYYSGTTPHPLNFTAP 1440
QY 1441 GSGHQGPFTGACGKSMSSVSLMGGRGVPLVYDRNHVTGASSSSSSSTKATLYPPLNP 1500
DB 1441 GSGHQGPFTGACGKSMSSVSLMGGRGVPLVYDRNHVTGASSSSSSSTKATLYPPLNP 1500
QY 1501 PPSPATDPSLYNMDMYSSNIPATARPYPYIIIRGMAPPPTTCTDVCDSYASARWKAS 1560
DB 1501 PPSPATDPSLYNMDMYSSNIPATARPYPYIIIRGMAPPPTTCTDVCDSYASARWKAS 1560
QY 1561 KYIYDLNDSDPYPPPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
DB 1561 KYIYDLNDSDPYPPPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 4
US-09-543-771B-3
; Sequence 3, Application US/09543771B
; Patent No. 6780609
; GENERAL INFORMATION:
; APPLICANT: Carulli, John P.
; APPLICANT: Little, Randall D.
; APPLICANT: Recker, Robert L.
; APPLICANT: Johnson, Mark L.
; TITLE OF INVENTION: High bone mass gene of 11q13.3
; FILE REFERENCE: 032796-014
; CURRENT APPLICATION NUMBER: US/09/543,771B
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
```

;
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 641
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-543-771B-3

Query Match 99.7%; Score 8715; DB 4; Length 1615;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1610; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MEAAPPGPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRVLDAGGKLESTIVS 60
DB 1 MEAAPPGPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRVLDAGGKLESTIVS 60

QY 61 GLEDAAAVDFQSKGAVYTWDSSEAIKQTYLNOTGAQVNVVLSGLVSPDGLACDWYVK 120
DB 61 GLEDAAAVDFQSKGAVYTWDSSEAIKQTYLNOTGAQVNVVLSGLVSPDGLACDWYVK 120

QY 121 KLYWTDSTNRIEVANLNGTSRKVLFWQDLQDPAIALDPAHGYMYTWDGETPRIERAG 180
DB 121 KLYWTDSTNRIEVANLNGTSRKVLFWQDLQDPAIALDPAHGYMYTWDGETPRIERAG 180

QY 181 MDGSTRKLIIVSDIYWPNGLTIDLEOKLYWADAKLSFIHRANLDGSPROKVVGSLTLP 240
DB 181 MDGSTRKLIIVSDIYWPNGLTIDLEOKLYWADAKLSFIHRANLDGSPROKVVGSLTLP 240

QY 241 FALTLSGDTLYWTDQWTRSHACNKRKTCGKKEILSALYSFMDIQVLSQERQPPFHTRCE 300
DB 241 FALTLSGDTLYWTDQWTRSHACNKRKTCGKKEILSALYSFMDIQVLSQERQPPFHTRCE 300

QY 301 EDNGGCSHLCLLSSEPPYTCACPTGVQVQDNGRTCKAGAEVLLARRTDLRRISLDP 360
DB 301 EDNGGCSHLCLLSSEPPYTCACPTGVQVQDNGRTCKAGAEVLLARRTDLRRISLDP 360

QY 361 DFTDIVLQVDDIRHAIAIDYDPLEGYVYTWDEVRATIRAYLDGSGAOTLVNTEINDPDG 420
DB 361 DFTDIVLQVDDIRHAIAIDYDPLEGYVYTWDEVRATIRAYLDGSGAOTLVNTEINDPDG 420

QY 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALHPVGMGLMYTWDMGE 480
DB 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALHPVGMGLMYTWDMGE 480

QY 481 NPKIECANLDGOERRVLYNASIGWPNGLALDQEGKYWDAKTDKIEVINVDGTKRRTL 540
DB 481 NPKIECANLDGOERRVLYNASIGWPNGLALDQEGKYWDAKTDKIEVINVDGTKRRTL 540

QY 541 LEDKLPHIFGFTLIGDFFIYWTQWRRSIRVHKVKASRDVLIIDQLPDLMLKAVNAKVW 600
DB 541 LEDKLPHIFGFTLIGDFFIYWTQWRRSIRVHKVKASRDVLIIDQLPDLMLKAVNAKVW 600

QY 601 GTNPCADRNGGCSHLCTFTPHATRCGCPGLELLSDMKTICVPEAFVFTSRAAHRISL 660
DB 601 GTNPCADRNGGCSHLCTFTPHATRCGCPGLELLSDMKTICVPEAFVFTSRAAHRISL 660

QY 661 ETNNNDVAIPITGVKEASALDFVSNNNHYWTDVSLKTIISRAFMNGSSVEHVFGLDYP 720
DB 661 ETNNNDVAIPITGVKEASALDFVSNNNHYWTDVSLKTIISRAFMNGSSVEHVFGLDYP 720

QY 721 EGMADVWGNLYWADTGTNRIEVARLDGQFVILVWRDLNPRSLALDPTKGYIYTWTEW 780
DB 721 EGMADVWGNLYWADTGTNRIEVARLDGQFVILVWRDLNPRSLALDPTKGYIYTWTEW 780

QY 781 GKGKPIVRAFMDGTNCMTLVKVGRLDITDADQRLYTDLDTNMISSNMLGOERVV 840
DB 781 GKGKPIVRAFMDGTNCMTLVKVGRLDITDADQRLYTDLDTNMISSNMLGOERVV 840

QY 841 IADDLPHFPGLTQYSDYIYWTDMNLHSIERADKTSGRNRTLIQGHLDPFVMDILVFHSSRQ 900
DB 841 IADDLPHFPGLTQYSDYIYWTDMNLHSIERADKTSGRNRTLIQGHLDPFVMDILVFHSSRQ 900

RESULT 5
US-09-544-398B-4
; Sequence 4, Application US/09544398B
; Patent No. 6770461
; GENERAL INFORMATION:
; APPLICANT: Carulli, John P.
; APPLICANT: Little, Randall D.
; APPLICANT: Recker, Robert R.
; APPLICANT: Johnson, Mark L.
; TITLE OF INVENTION: High bone mass gene of 11q13.3
; FILE REFERENCE: 032796-013
; CURRENT APPLICATION NUMBER: US/09/544,398B
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; NUMBER OF SEQ ID NOS: 641
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4

```
; LENGTH: 1615
; TYPE: PRX
; ORGANISM: Homo sapiens
US-09-544-398B-4

Query Match          99.6%; Score 8706; DB 4; Length 1615;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1609; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MEAAPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRIVLDAGGVKLESTIVVS 60
DB 1 MEAAPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRIVLDAGGVKLESTIVVS 60

QY 61 GLEDAAAVDFQSKGAVYWTVDSEAIKQTYLNOTGAQVQNVVVISGLVSPDGLACDWYVK 120
DB 61 GLEDAAAVDFQSKGAVYWTVDSEAIKQTYLNOTGAQVQNVVVISGLVSPDGLACDWYVK 120

QY 121 KLYWTDSETNRIEVANLNGTSKVLFWQDLDOPRAIALDPAGHYMYWTDWGETPRIERAG 180
DB 121 KLYWTDSETNRIEVANLNGTSKVLFWQDLDOPRAIALDPAGHYMYWTDWGETPRIERAG 180

QY 181 MDGSTRKIIVSDIYWPNGLTIDLEOKLYWADAKLSPIHRANLDGSPROKVEGSLTHP 240
DB 181 MDGSTRKIIVSDIYWPNGLTIDLEOKLYWADAKLSPIHRANLDGSPROKVEGSLTHP 240

QY 241 FALTLSGDTLYWTDWQTSIIHACNKRITGKKEILSALYSPMDIQVLQOERQPPFHTRCE 300
DB 241 FALTLSGDTLYWTDWQTSIIHACNKRITGKKEILSALYSPMDIQVLQOERQPPFHTRCE 300

QY 301 EDNGGCSHLCLLSPSEPFYTCACPTGVOLQNGRTCKAGABEVLLARRTDLRRLSLDTP 360
DB 301 EDNGGCSHLCLLSPSEPFYTCACPTGVOLQNGRTCKAGABEVLLARRTDLRRLSLDTP 360

QY 361 DFTDVLQVDDIRHAIADYDPLEGYVWTDDEVRAIRRAYLDGSGAQLVNTENDPDG 420
DB 361 DFTDVLQVDDIRHAIADYDPLEGYVWTDDEVRAIRRAYLDGSGAQLVNTENDPDG 420

QY 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKLVSEDLDEPRALHVPVGLMYWTDWGE 480
DB 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKLVSEDLDEPRALHVPVGLMYWTDWGE 480

QY 481 NPKIECANLDQERRVLVNASLGNWENGLALDQEGKLYWDAKTDKIEVINVDGTKRTL 540
DB 481 NPKIECANLDQERRVLVNASLGNWENGLALDQEGKLYWDAKTDKIEVINVDGTKRTL 540

QY 541 LEDKLPHIFGTLGDFIYTWDMQRRSIEVHKVKAERDVIIQDLPDLMLGLKAVNVAKV 600
DB 541 LEDKLPHIFGTLGDFIYTWDMQRRSIEVHKVKAERDVIIQDLPDLMLGLKAVNVAKV 600

QY 601 GTNPCADNRGCGSHLCFTPHATRCGCPIGLELLSDMKTCIVPEAFIVFTSRAATHRISL 660
DB 601 GTNPCADNRGCGSHLCFTPHATRCGCPIGLELLSDMKTCIVPEAFIVFTSRAATHRISL 660

QY 661 ETNNNDVAIPLTGKVASALDFVSNHHIYWTDVSLKTIISRAFMNGSSVEHVVFGLDYP 720
DB 661 ETNNNDVAIPLTGKVASALDFVSNHHIYWTDVSLKTIISRAFMNGSSVEHVVFGLDYP 720

QY 721 EGMVDMWNGKLYWADTGTNRIEVARLDQPROVLVWRDLNPRSLALDPTKGYIYTWTEW 780
DB 721 EGMVDMWNGKLYWADTGTNRIEVARLDQPROVLVWRDLNPRSLALDPTKGYIYTWTEW 780

QY 781 GGPRIVRAFMDGTNCMTLVKVGRLDITDYADORLYWTDLTNNMIESNNMLQGERVV 840
DB 781 GGPRIVRAFMDGTNCMTLVKVGRLDITDYADORLYWTDLTNNMIESNNMLQGERVV 840

QY 841 IADDLPHFPGLTQYSDIYWTDWNLHLSIERADKTSGRNRTLIQGHLDFFWMDILVFHSSRQ 900
DB 841 IADDLPHFPGLTQYSDIYWTDWNLHLSIERADKTSGRNRTLIQGHLDFFWMDILVFHSSRQ 900

QY 901 DGLNDGMHNNCGGOLCLAIIPGHRCCGASHYTLDPSSRNCSPPTTLLFSQKSAISRMI 960
DB 901 DGLNDGMHNNCGGOLCLAIIPGHRCCGASHYTLDPSSRNCSPPTTLLFSQKSAISRMI 960

QY 961 PDQHSPLDILPLHCLRNKVIDYDPLDKFYWVDGRQNIKRAKXDDGTQPPVLTSLSQG 1020
DB 961 PDQHSPLDILPLHCLRNKVIDYDPLDKFYWVDGRQNIKRAKXDDGTQPPVLTSLSQG 1020

QY 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIIVVNAERG 1080
DB 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIIVVNAERG 1080

QY 1081 LYFTNQMDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESC 1140
DB 1081 LYFTNQMDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESC 1140

QY 1141 LSGANRLTLEDAIVQPLGLTILGKHLIYDROOQMIERVEKTTGDKRTRIQGRVAHLTG 1200
DB 1141 LSGANRLTLEDAIVQPLGLTILGKHLIYDROOQMIERVEKTTGDKRTRIQGRVAHLTG 1200

QY 1201 IHAVEVSLBESFAHPACARDNGGCSHICIAKGDGTPRCSCPVHLVLLQNLITCGEPPTCS 1260
DB 1201 IHAVEVSLBESFAHPACARDNGGCSHICIAKGDGTPRCSCPVHLVLLQNLITCGEPPTCS 1260

QY 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEGCPVCSAAQFPCARGQCVDLRLRCDGE 1320
DB 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEGCPVCSAAQFPCARGQCVDLRLRCDGE 1320

QY 1321 ADCODRDEADCDALCLPNQFRCSGOCVLIKQOCDFPDICDGSDELWCEITKPPSDDS 1380
DB 1321 ADCODRDEADCDALCLPNQFRCSGOCVLIKQOCDFPDICDGSDELWCEITKPPSDDS 1380

QY 1381 PAHSAAGPVGIIILSLFVMGGVYFVQRCVQRYAGANGPPHYVSGTTPHVPINFIAP 1440
DB 1381 PAHSAAGPVGIIILSLFVMGGVYFVQRCVQRYAGANGPPHYVSGTTPHVPINFIAP 1440

QY 1441 GGSQHGPTGTIACGKSMSSVSLMGRGGVPLYDRNHVHTGASSSSSSSTKATLPPILNP 1500
DB 1441 GGSQHGPTGTIACGKSMSSVSLMGRGGVPLYDRNHVHTGASSSSSSSTKATLPPILNP 1500

QY 1501 PPSPATPSLYNDMFYSSNIPATAPVPIYIRGMAPPPTPCSTDVCDSDYSASRWKAS 1560
DB 1501 PPSPATPSLYNDMFYSSNIPATAPVPIYIRGMAPPPTPCSTDVCDSDYSASRWKAS 1560

QY 1561 KYLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
DB 1561 KYLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 6
US-09-543-771B-4
; Sequence 4, Application US/09543771B
; Patent No. 6780609
; GENERAL INFORMATION:
; APPLICANT: Carulli, John P.
; APPLICANT: Little, Randall D.
; APPLICANT: Recker, Robert R.
; APPLICANT: Johnson, Mark L.
; TITLE OF INVENTION: High bone mass gene of 11q13.3
; CURRENT APPLICATION NUMBER: US/09/543,771B
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 641
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1615
; TYPE: PRX
; ORGANISM: Homo sapiens
US-09-543-771B-4

Query Match          99.6%; Score 8706; DB 4; Length 1615;
```


Best Local Similarity 99.6%; Pred. No. 0;			
Matches 1609; Conservative 2; Mismatches 4; Indels 0; Gaps 0;			
Qy	1	MEAAPGPPWLLLLLLLLLLLALCCGAPAPAAAPLPLPANRRDVRVLVDAGGKLESTIVS	60
Db	1	MEAAPGPPWLLLLLLLLLLLALCCGAPAPAAAPLPLPANRRDVRVLVDAGGKLESTIVS	60
Qy	61	GLEDAAAVDFGSKGAVTWDVSEBAIKQTYLNQTAAGVQNVVWISGLVSPDGLACDWGK	120
Db	61	GLEDAAAVDFGSKGAVTWDVSEBAIKQTYLNQTAAGVQNVVWISGLVSPDGLACDWGK	120
Qy	121	KLYTDSNTNRIEVANLNGTSRKVLFWQDLOPRAIALDPAGHYMYTWDWGETPRIERAG	180
Db	121	KLYTDSNTNRIEVANLNGTSRKVLFWQDLOPRAIALDPAGHYMYTWDWGETPRIERAG	180
Qy	181	MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSPIHRANLDGSPROKRVGSLTTP	240
Db	181	MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSPIHRANLDGSPROKRVGSLTTP	240
Qy	241	FALTLSGDTLYTWDQTRSIIHACNKRTGKKEIILSALYSPMDIQVLSQERQPFPHTRCE	300
Db	241	FALTLSGDTLYTWDQTRSIIHACNKRTGKKEIILSALYSPMDIQVLSQERQPFPHTRCE	300
Qy	301	EDNGGCSHLCLSSEPPYTACPTGVOLQNGRTCKAGABEVLLLAARTDLRISLDT	360
Db	301	EDNGGCSHLCLSSEPPYTACPTGVOLQNGRTCKAGABEVLLLAARTDLRISLDT	360
Qy	361	DFTDILVQVDDIRHAIAIDYDPLGYYVWTDDEVRATRRAYLDGSGAQLVNTNINPDG	420
Db	361	DFTDILVQVDDIRHAIAIDYDPLGYYVWTDDEVRATRRAYLDGSGAQLVNTNINPDG	420
Qy	421	IADVARNLYTWTGTDIRIEVTRNGTSRKILVSIEDLDEPRALALHPVWGLMTWTDGE	480
Db	421	IADVARNLYTWTGTDIRIEVTRNGTSRKILVSIEDLDEPRALALHPVWGLMTWTDGE	480
Qy	481	NPKIECANLDQERRVLVNASLGNPGLALDQEGKLYWDAKTKIEVINVDGTKRRTL	540
Db	481	NPKIECANLDQERRVLVNASLGNPGLALDQEGKLYWDAKTKIEVINVDGTKRRTL	540
Qy	541	LEDKLPHIFGTLGDFRYTWDQRRSIRVHKVKAERDVIIDQIPDLMLGKANYAKV	600
Db	541	LEDKLPHIFGTLGDFRYTWDQRRSIRVHKVKAERDVIIDQIPDLMLGKANYAKV	600
Qy	601	GTNFCADRNGCSHLCPFTPHATRCGCPIGLELLSDMKTICVPEAFVFTSRAAHRISL	660
Db	601	GTNFCADRNGCSHLCPFTPHATRCGCPIGLELLSDMKTICVPEAFVFTSRAAHRISL	660
Qy	661	ETNNNDVAIPUTGVKEASALDFVSNHHIYWTDSVSLKTIISAFWNGSSVEHVFEGLDYP	720
Db	661	ETNNNDVAIPUTGVKEASALDFVSNHHIYWTDSVSLKTIISAFWNGSSVEHVFEGLDYP	720
Qy	721	EGMAVDWNGKLYWADTGTNRIEVARLDGQPROVLVWRDLNPRSLALDPTKGYIYTEW	780
Db	721	EGMAVDWNGKLYWADTGTNRIEVARLDGQPROVLVWRDLNPRSLALDPTKGYIYTEW	780
Qy	781	GCKPRIVRAFMGNTNCTLVKVRANDLTIDYADQLYWTDLTNMIESNNMLGQERVV	840
Db	781	GCKPRIVRAFMGNTNCTLVKVRANDLTIDYADQLYWTDLTNMIESNNMLGQERVV	840
Qy	841	IADDLPHFPGLTQYSDVIYWTDMNLHSIERADKTSGRNRTLIOGHLDPFVMDILVFHSRQ	900
Db	841	IADDLPHFPGLTQYSDVIYWTDMNLHSIERADKTSGRNRTLIOGHLDPFVMDILVFHSRQ	900
Qy	901	DGLNDCHNNGCQCQLCLAIIPGHRCCGASHYTLDPSSRNCSPPTTLLFSQKSAISRM	960
Db	901	DGLNDCHNNGCQCQLCLAIIPGHRCCGASHYTLDPSSRNCSPPTTLLFSQKSAISRM	960
Qy	961	PDDQHSPLIILPLHGLRNKAIIDYDPLDKFYWVDGRQNIKRAKDGTQPPVLTSQGG	1020
Db	961	PDDQHSPLIILPLHGLRNKAIIDYDPLDKFYWVDGRQNIKRAKDGTQPPVLTSQGG	1020
Qy	1021	NPDRQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG	1080
Db	1021	NPDRQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG	1080

Db	1021	NPDRQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG	1080
Qy	1081	LYFTNMODRAAKIERAALDGTREVLFTTGLIRVALVVDNTLGLFWVDADLKRISSCD	1140
Db	1081	LYFTNMODRAAKIERAALDGTREVLFTTGLIRVALVVDNTLGLFWVDADLKRISSCD	1140
Qy	1141	LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIERVEKTTGDKETRIQGRVAHLTG	1200
Db	1141	LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIERVEKTTGDKETRIQGRVAHLTG	1200
Qy	1201	IHAVEVSLBFSAPCARNDGGCSHICIAKGDGTPRCSPVHLVLQNLTCGEPPTCS	1260
Db	1201	IHAVEVSLBFSAPCARNDGGCSHICIAKGDGTPRCSPVHLVLQNLTCGEPPTCS	1260
Qy	1261	PDQFACATGEIDCIPGAWRCDGPFECDDQSDDEECPCVCSAAQFPCCARGQCVDLRLRCDGE	1320
Db	1261	PDQFACATGEIDCIPGAWRCDGPFECDDQSDDEECPCVCSAAQFPCCARGQCVDLRLRCDGE	1320
Qy	1321	ADCDRSDDEADCDACILPNQFRCASGQCVLIKQCCDSFPDCIDGSDDELWCBITKPPSDDS	1380
Db	1321	ADCDRSDDEADCDACILPNQFRCASGQCVLIKQCCDSFPDCIDGSDDELWCBITKPPSDDS	1380
Qy	1381	PAHSSAIGPVIGIILSLFVWGVYFVQVQVQRYAGANGPPHYYVSGTTPHVLNFIAP	1440
Db	1381	PAHSSAIGPVIGIILSLFVWGVYFVQVQVQRYAGANGPPHYYVSGTTPHVLNFIAP	1440
Qy	1441	GGSHQGPFTGIACGKSMSSVSLMGGRGVPLYDRNHVGTGASSSSSKATLYPPIILNP	1500
Db	1441	GGSHQGPFTGIACGKSMSSVSLMGGRGVPLYDRNHVGTGASSSSSKATLYPPIILNP	1500
Qy	1501	PPSPATDPSLYNMDMFYSSNIPATAPRYPIYIRGMAPPTTTCSTDVCDSDYSASRWKAS	1560
Db	1501	PPSPATDPSLYNMDMFYSSNIPATAPRYPIYIRGMAPPTTTCSTDVCDSDYSASRWKAS	1560
Qy	1561	KYIYDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS	1615
Db	1561	KYIYDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS	1615

RESULT 7

US-09-060-299-4
; Sequence 4, Application US/09060299
; Patent No. 6545137
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; APPLICANT: Hess, John W
; APPLICANT: Caskey, Charles T
; APPLICANT: Cox, Roger D
; APPLICANT: Gerhold, David
; APPLICANT: Hammond, Holly
; APPLICANT: Hey, Patricia
; APPLICANT: Kawaguchi, Yoshihiko
; APPLICANT: Merriman, Tony R
; APPLICANT: Metzger, Michael L
; TITLE OF INVENTION: No. 6545137el Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6545137th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,299
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/043,553

; FILING DATE: 15-APR-1997

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 60/048,740

; FILING DATE: 05-JUN-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: B.J.Sadoff

; REGISTRATION NUMBER: 36,663

; REFERENCE/DOCKET NUMBER: 620-35

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703)816-4091

; TELEFAX: (703)816-4100

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1591 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

US-09-060-299-4

Query Match 98.5%; Score 8607; DB 4; Length 1591;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 25 CPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQPSKGAIVYTDVSE 84
DB 1 CPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQPSKGAIVYTDVSE 60

QY 85 EAIKQTYLNQGAQVNVVIVSLVSPDGLADWVGKLYWTDSETNRIEVANLNGTSRKV 144
DB 61 EAIKQTYLNQGAQVNVVIVSLVSPDGLADWVGKLYWTDSETNRIEVANLNGTSRKV 120

QY 145 LFQWQDLPRAIALDPAGHYMTWDTGETPRIERAGMDGSTRKIIIVDSDIYWPNGLTIDL 204
DB 121 LFQWQDLPRAIALDPAGHYMTWDTGETPRIERAGMDGSTRKIIIVDSDIYWPNGLTIDL 180

QY 205 EEQKLYWADAKLSFTHRANLDGSRQKVEGSLTHPFAITLSGDTLYWTDQTRSIIHACN 264
DB 181 EEQKLYWADAKLSFTHRANLDGSRQKVEGSLTHPFAITLSGDTLYWTDQTRSIIHACN 240

QY 265 KRTGGRKEIISALYSPMDIQVLQOERQPPFTHRCEDNGGCSHLCLLSPSEPFYTACP 324
DB 241 KRTGGRKEIISALYSPMDIQVLQOERQPPFTHRCEDNGGCSHLCLLSPSEPFYTACP 300

QY 325 TGVQLQNGRTCKGAEEVLLARLTDLRISLTDPTDFTDVLQVDDIRHAIAIDYDPLE 384
DB 301 TGVQLQNGRTCKGAEEVLLARLTDLRISLTDPTDFTDVLQVDDIRHAIAIDYDPLE 360

QY 385 GYVYVTDDEVAIRRAYLDGSGAQLVNTENIDPDGIADVAVARNLYWTDGTDRIEVTR 444
DB 361 GYVYVTDDEVAIRRAYLDGSGAQLVNTENIDPDGIADVAVARNLYWTDGTDRIEVTR 420

QY 445 LNSTSRKILVSEDLDEPRALHVPVGLMYWTDGENPKIECANLDGQERRVLYNASLIGW 504
DB 421 LNSTSRKILVSEDLDEPRALHVPVGLMYWTDGENPKIECANLDGQERRVLYNASLIGW 480

QY 505 PNGLALDLQEGKLYWGDAKTDKIEVINVDGTRKRTLLDKLPHIFGFTLLGDFIYTDWQ 564
DB 481 PNGLALDLQEGKLYWGDAKTDKIEVINVDGTRKRTLLDKLPHIFGFTLLGDFIYTDWQ 540

QY 565 RRSIERVHKVAKSRDVIIDQLPDLMLGLKANVAVKVGTPNCPADRNGGCSHLCPFTPHATR 624
DB 541 RRSIERVHKVAKSRDVIIDQLPDLMLGLKANVAVKVGTPNCPADRNGGCSHLCPFTPHATR 600

QY 625 CGCPIGLELSDMTKCIIVPEAFVFTSRAATHRSLETNNDVVAIPITGVKEASALDPDV 684
DB 601 CGCPIGLELSDMTKCIIVPEAFVFTSRAATHRSLETNNDVVAIPITGVKEASALDPDV 660

QY 685 SNNHIYWDVSLKTIISRAFMNGSSVHVVEFGLDYPEGMAVDMGKNLYWADTGTNRIEV 744
DB 661 SNNHIYWDVSLKTIISRAFMNGSSVHVVEFGLDYPEGMAVDMGKNLYWADTGTNRIEV 720

QY 745 ARLDGQFQVLYWDLNPNRSLALDPTKGYIYWTWEGGKPRIVAFMDGTNCMTLVDKVG 804
DB 721 ARLDGQFQVLYWDLNPNRSLALDPTKGYIYWTWEGGKPRIVAFMDGTNCMTLVDKVG 780

QY 805 RANDLTIDYADQRLYWTDLDTNMIESSNMLQGERVVIADDLPHFPGLTQYSDIYIYWTOWN 864
DB 781 RANDLTIDYADQRLYWTDLDTNMIESSNMLQGERVVIADDLPHFPGLTQYSDIYIYWTOWN 840

QY 865 LHSIERADKTSGRNRTLQGHLDLFMDILVPHSSRODGLNDCMHNNGCGQLCLAIPOGH 924
DB 841 LHSIERADKTSGRNRTLQGHLDLFMDILVPHSSRODGLNDCMHNNGCGQLCLAIPOGH 900

QY 925 RCGCASHYTLDPSSRNCSPPFTTFLFSOKSAISMIIPDOQHSPLILPLHGRNVKADY 984
DB 901 RCGCASHYTLDPSSRNCSPPFTTFLFSOKSAISMIIPDOQHSPLILPLHGRNVKADY 960

QY 985 DPLDKFIYWDGRQNIKRAKDDGTQPFVLTSLSQQNPDRQPHDLSIDYISRTLPWTCEA 1044
DB 961 DPLDKFIYWDGRQNIKRAKDDGTQPFVLTSLSQQNPDRQPHDLSIDYISRTLPWTCEA 1020

QY 1045 TTNINVHRLSGEAMGVLRGRDKPRAIIVNAERGILYFTNMQDRAAKIERAALDGTRE 1104
DB 1021 TTNINVHRLSGEAMGVLRGRDKPRAIIVNAERGILYFTNMQDRAAKIERAALDGTRE 1080

QY 1105 VLFTTGLIRPVALVVDNTLGLKFWVDADLKRIESCDLSGANRLTLEDANIYQPLGLTILG 1164
DB 1081 VLFTTGLIRPVALVVDNTLGLKFWVDADLKRIESCDLSGANRLTLEDANIYQPLGLTILG 1140

QY 1165 KHLWIDRQOQMIERVEKTTGDKRTRIQRVAHLTGIHAEVEVLEEFSAHPCARDNGGC 1224
DB 1141 KHLWIDRQOQMIERVEKTTGDKRTRIQRVAHLTGIHAEVEVLEEFSAHPCARDNGGC 1200

QY 1225 SHICIAKGDGTPRCSCPVLVLLQNLATCGEPPTCSPOFACATGEIDCIPGANRCDFP 1284
DB 1201 SHICIAKGDGTPRCSCPVLVLLQNLATCGEPPTCSPOFACATGEIDCIPGANRCDFP 1260

QY 1285 ECDQSDDEEGCPVCSAAQFPCARQCQVDLRRCDEADCDRSDDEADCDALCLPNQFRCA 1344
DB 1261 ECDQSDDEEGCPVCSAAQFPCARQCQVDLRRCDEADCDRSDDEADCDALCLPNQFRCA 1320

QY 1345 SGQCVLTKQCCDPPDCIDGSDLELMCEITKPPSDSDSAHSAISGIVTIGIILSLFVMGCV 1404
DB 1321 SGQCVLTKQCCDPPDCIDGSDLELMCEITKPPSDSDSAHSAISGIVTIGIILSLFVMGCV 1380

QY 1405 FVQORVVCORVAGANGFPFHEYVSGTVPVPLNFTAPGSOHGPFTGTACGSKMSSVSLM 1464
DB 1381 FVQORVVCORVAGANGFPFHEYVSGTVPVPLNFTAPGSOHGPFTGTACGSKMSSVSLM 1440

QY 1465 GGRGGVLYDRNHVHTGASSSSSSSTKATLYPILNPPSPATDPSLYNMDMFYSSNIPAT 1524
DB 1441 GGRGGVLYDRNHVHTGASSSSSSSTKATLYPILNPPSPATDPSLYNMDMFYSSNIPAT 1500

QY 1525 ARPYRPIIRGMAPPPTTPTCSTDVCDSDYSASRWKASKYIYDLNDSDDYPPPPPTPHSOYL 1584
DB 1501 VRPYRPIIRGMAPPPTTPTCSTDVCDSDYSASRWKASKYIYDLNDSDDYPPPPPTPHSOYL 1560

QY 1585 SAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
DB 1561 SAEDSCPPSPATERSYFHLPPPPSPCTDSS 1591

RESULT 8

US-09-060-299-43

; Sequence 43, Application US/09060299

; Patent No. 6545137

; GENERAL INFORMATION:

; APPLICANT: Todd, John A

; APPLICANT: Hess, John W

; APPLICANT: Caskey, Charles T

; APPLICANT: Cox, Roger D

; APPLICANT: Gerhold, David

; APPLICANT: Hammond, Holly

; APPLICANT: Hey, Patricia

; APPLICANT: Kawaguchi, Yoshihiko

; APPLICANT: Merriman, Tony R

; APPLICANT: Metzker, Michael L

TITLE OF INVENTION: No. 6545137e1 Receptor
NUMBER OF SEQUENCES: 455
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon and Vanderhye
STREET: 1100 No. 6545137th Glebe Road, Eighth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: US
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/060,299
FILING DATE: 15-APR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/043,553
FILING DATE: 15-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/048,740
FILING DATE: 05-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: B.J.Sadoff
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 620-35
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4091
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 1591 amino acids
TYPE: amino acid
TOPOLOGY: linear
us-09-060-299-43

Query Match 98.5%; Score 8607; DB 4; Length 1591;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	25	CPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQFSKGVYWTDVSE	84
Db	1	CPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQFSKGVYWTDVSE	60
Qy	85	BAIKQTYLNOTGAQVONVVISGLVSPDGLACDWGKLYWTDSTNRIEVANLNGTSRKV	144
Db	61	BAIKQTYLNOTGAQVONVVISGLVSPDGLACDWGKLYWTDSTNRIEVANLNGTSRKV	120
Qy	145	LFWQDLQPRALDPAHGYWYTDWGETPRIERAGMDGSTRKIIIVDSDIYWPNGLTIDL	204
Db	121	LFWQDLQPRALDPAHGYWYTDWGETPRIERAGMDGSTRKIIIVDSDIYWPNGLTIDL	180
Qy	205	EEQKLYWADAKLSFIHRANLDGSPQKQVVEGSLTHFPALTLSGDTLYWTDQTRSIIHACN	264
Db	181	EEQKLYWADAKLSFIHRANLDGSPQKQVVEGSLTHFPALTLSGDTLYWTDQTRSIIHACN	240
Qy	265	KRTGKKEKILSALYSPMDIOVLQERQPPFTHRCEDNGGCSHCLLSPSPEPYTCACP	324
Db	241	KRTGKKEKILSALYSPMDIOVLQERQPPFTHRCEDNGGCSHCLLSPSPEPYTCACP	300
Qy	325	TGVQLQNGRTCKAGAEVLLARLTDLRISLSDTPDFTDVLQVDDIRHAIADYDPLE	384
Db	301	TGVQLQNGRTCKAGAEVLLARLTDLRISLSDTPDFTDVLQVDDIRHAIADYDPLE	360
Qy	385	GYVYWTDDVEAIRRAVLDGSGAQLVNTENIDPDGIADVAVARNLYWTDGTDRIEVTR	444
Db	361	GYVYWTDDVEAIRRAVLDGSGAQLVNTENIDPDGIADVAVARNLYWTDGTDRIEVTR	420
Qy	445	LNQTSRKILVSEDDLEPPRAIALHPVWMLYWTDMGENPKIECANLDGQERRVLVNASLIGW	504
Db	421	LNQTSRKILVSEDDLEPPRAIALHPVWMLYWTDMGENPKIECANLDGQERRVLVNASLIGW	480

Qy	505	PNGIALDLQEGKLYWGDAKTDKIEVINVDGTRKRTLLDKLPHIFGFTLLGDFIYTDWQ	564
Db	481	PNGIALDLQEGKLYWGDAKTDKIEVINVDGTRKRTLLDKLPHIFGFTLLGDFIYTDWQ	540
Qy	565	RRSIERVHKVAKSRDVIIDQPLDMLGKLVANVAVKVTGTPCADRNGGCSHLCFFTPHATR	624
Db	541	RRSIERVHKVAKSRDVIIDQPLDMLGKLVANVAVKVTGTPCADRNGGCSHLCFFTPHATR	600
Qy	625	CGCPGLELLSDMKTCIVPRAFLVFTSRAAHRISLETNNNDVAIPLTGVEASALDFDV	684
Db	601	CGCPGLELLSDMKTCIVPRAFLVFTSRAAHRISLETNNNDVAIPLTGVEASALDFDV	660
Qy	685	SNNHYYTVDVSLKTIISRAFMNGSSVEHVVEFGLDYPGMAVDMGKNLYWADTGTNRLEV	744
Db	661	SNNHYYTVDVSLKTIISRAFMNGSSVEHVVEFGLDYPGMAVDMGKNLYWADTGTNRLEV	720
Qy	745	ARLDGQFQVLRVWRDLNPRSLALDPTKGYIYTWENGKPKRIVRAFMDGTNCMLTLDKVG	804
Db	721	ARLDGQFQVLRVWRDLNPRSLALDPTKGYIYTWENGKPKRIVRAFMDGTNCMLTLDKVG	780
Qy	805	RANDLTIDYADORLYWTDLTNMISSNMLGQBRVVIADDLPHFPGLTOYSDIYTWTDN	864
Db	781	RANDLTIDYADORLYWTDLTNMISSNMLGQBRVVIADDLPHFPGLTOYSDIYTWTDN	840
Qy	865	LHSIERADKTSGRNRTLIQGHLDFFVMDILVFHSSRQDGLNDCMENNCGCQLCLAIPEGH	924
Db	841	LHSIERADKTSGRNRTLIQGHLDFFVMDILVFHSSRQDGLNDCMENNCGCQLCLAIPEGH	900
Qy	925	RCGCASHYTLDPSSRNCSPPTFLPSQKSAISMIIPDDQHSPLDILPLHGLRNVAIDY	984
Db	901	RCGCASHYTLDPSSRNCSPPTFLPSQKSAISMIIPDDQHSPLDILPLHGLRNVAIDY	960
Qy	985	DPLDKFIYWDGRONIKRAKDDGTOPFVLTSLSGQNPDRPHDLSIDYISRTLPWTCEA	1044
Db	961	DPLDKFIYWDGRONIKRAKDDGTOPFVLTSLSGQNPDRPHDLSIDYISRTLPWTCEA	1020
Qy	1045	TNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERGILYFTNMQDRAAKIERAALDGTRE	1104
Db	1021	TNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERGILYFTNMQDRAAKIERAALDGTRE	1080
Qy	1105	VLFTTGLIRPVALVVDNTLQKLFWDADLKRIBSCDLSGANRLTLEDANIYVQPLGLTILG	1164
Db	1081	VLFTTGLIRPVALVVDNTLQKLFWDADLKRIBSCDLSGANRLTLEDANIYVQPLGLTILG	1140
Qy	1165	KHLYWIDRQOQMIERVEKTTGDKETRIQGRVAHLTGTHAVEEVSLEBFSAPPCARDNGCC	1224
Db	1141	KHLYWIDRQOQMIERVEKTTGDKETRIQGRVAHLTGTHAVEEVSLEBFSAPPCARDNGCC	1200
Qy	1225	SHICIAKGDGTPRCSCPVLVLLQNLITCGEPPTCSPDQFACATGEIDCIPGAWRCDFGP	1284
Db	1201	SHICIAKGDGTPRCSCPVLVLLQNLITCGEPPTCSPDQFACATGEIDCIPGAWRCDFGP	1260
Qy	1285	ECDQSDDEEGCPVCSAAQFPFCARGQCVDLRLRCDEADCDQSDDEADCDAI CLNPQFRCA	1344
Db	1261	ECDQSDDEEGCPVCSAAQFPFCARGQCVDLRLRCDEADCDQSDDEADCDAI CLNPQFRCA	1320
Qy	1345	SGQCVLJKQCDSPDCIDGSDDELWCEITKPPSDSDSAHSSAIGPVIGIILSLFVWGGVY	1404
Db	1321	SGQCVLJKQCDSPDCIDGSDDELWCEITKPPSDSDSAHSSAIGPVIGIILSLFVWGGVY	1380
Qy	1405	FVQORVVCQYAGANGFPFHEYVSGTTPHVPLNFTAPGSGHQHPTGIACGKSMSSVSLM	1464
Db	1381	FVQORVVCQYAGANGFPFHEYVSGTTPHVPLNFTAPGSGHQHPTGIACGKSMSSVSLM	1440
Qy	1465	GGRGVPLYDRNHVGTGASSSSSTKATLYPPIINLPPSPATDPSLYNMDMFPYSNIPAT	1524
Db	1441	GGRGVPLYDRNHVGTGASSSSSTKATLYPPIINLPPSPATDPSLYNMDMFPYSNIPAT	1500
Qy	1525	ARPYRPYIIRGMAPPITPCSTDVCDSDYSASRWKASKYVLDLNSDSDPYPPPPPHSOYL	1584
Db	1501	VRPYRPYIIRGMAPPITPCSTDVCDSDYSASRWKASKYVLDLNSDSDPYPPPPPHSOYL	1560
Qy	1585	SAEDSCPPSPATERSYFHLPPPPSPCTDSS	1615

Db 1561 SAEDSCPPSPATERSYFHLFPFPPSPCTDSS 1591
|||||
RESULT 9
US-09-402-923A-4
; Sequence 4, Application US/09402923A
; Patent No. 6555654
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; Hess, John W
; Caskey, Charles T
; Cox, Roger D
; Gerhold, David
; Hammond, Holly
; Hey, Patricia
; Kawaguchi, Yoshihiko
; Merriman, Tony R
; Metzker, Michael L
; TITLE OF INVENTION: No. 6555654e1 LDL-Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6555654th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; ZIP: VA 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA: US/09/402,923A
; FILING DATE: 14-Feb-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB98/01102
; FILING DATE: 15-APR-1998
; APPLICATION NUMBER: US 60/043,553
; FILING DATE: 15-APR-1997
; APPLICATION NUMBER: US 60/048,740
; FILING DATE: 05-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: B.J.Sadoff
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 620-81
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4091
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1591 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-402-923A-4
Query Match 98.5%; Score 8607; DB 4; Length 1591;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 25 CPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQFSGAVVTDVSE 84
Db 1 CPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQFSGAVVTDVSE 60
Qy 85 EAIQTILNQTAQVNVVIVSGVSPDGLACDWVGVKKLYWTDSTNRIEVANLNGTSRKY 144
Db 61 EAIQTILNQTAQVNVVIVSGVSPDGLACDWVGVKKLYWTDSTNRIEVANLNGTSRKY 120
Qy 145 LFWDQDLPRAIALDPAGHYWYWDWGETPRIERAGMDGSTRKLIIVDSDIYWPNGLTIDL 204
Db 121 LFWDQDLPRAIALDPAGHYWYWDWGETPRIERAGMDGSTRKLIIVDSDIYWPNGLTIDL 180

Qy 205 EEQKLYWADAKLSFTHRANLDGSPQKQVVEGSLTHPPFALTLSGDTLYWTDQTRSIHACN 264
Db 181 EEQKLYWADAKLSFTHRANLDGSPQKQVVEGSLTHPPFALTLSGDTLYWTDQTRSIHACN 240
Qy 265 KRTGGRKEIISALYSPMDIOVLQSOEROPFPHTRCEDNGGCSHCLLSPEBPFTYCAP 324
Db 241 KRTGGRKEIISALYSPMDIOVLQSOEROPFPHTRCEDNGGCSHCLLSPEBPFTYCAP 300
Qy 325 TGVQLQDNMGRTCKAGAEVLLARRTDLRRISLQTPDFDVLQVDDIRHAIADYDPLE 384
Db 301 TGVQLQDNMGRTCKAGAEVLLARRTDLRRISLQTPDFDVLQVDDIRHAIADYDPLE 360
Qy 385 GYVYVTDDEVAIRRAYLDGSGAQTLLVNTENDPGIAVDMVARNLYWTDGTDRIEVTR 444
Db 361 GYVYVTDDEVAIRRAYLDGSGAQTLLVNTENDPGIAVDMVARNLYWTDGTDRIEVTR 420
Qy 445 LNSTSRKILVSEDLDEPAIAHPVMGLMYWTDGENPKIECANLDGQERRVIVNASLGM 504
Db 421 LNSTSRKILVSEDLDEPAIAHPVMGLMYWTDGENPKIECANLDGQERRVIVNASLGM 480
Qy 505 ENGLALDLOEGKLYWDAKTDKIEVINVDGTRKRTLLKPLHIFGFTLLGDFIYWTQW 564
Db 481 ENGLALDLOEGKLYWDAKTDKIEVINVDGTRKRTLLKPLHIFGFTLLGDFIYWTQW 540
Qy 565 RRSIERVHKVKAASRDVIIDQLPDLMLGLKAVNAVAVGTNPICADNRNGGCSHLCFFTPHATR 624
Db 541 RRSIERVHKVKAASRDVIIDQLPDLMLGLKAVNAVAVGTNPICADNRNGGCSHLCFFTPHATR 600
Qy 625 CGCFIGLELLSDMKTCIVPEAFVFTSRAAIIHRSLETNNNDVAIPLTGKVEASALDFDV 684
Db 601 CGCFIGLELLSDMKTCIVPEAFVFTSRAAIIHRSLETNNNDVAIPLTGKVEASALDFDV 660
Qy 685 SNNHYYWTDVSLKTIISRAFNMGSSVEHVVEFGLDYPEGMAVDMWGNKLYWADTGNRIEV 744
Db 661 SNNHYYWTDVSLKTIISRAFNMGSSVEHVVEFGLDYPEGMAVDMWGNKLYWADTGNRIEV 720
Qy 745 ARLDGQFQVILVWRDLNPRSLALDPTKGYIYTWENGKPRIVAFMDGTNCMTLVKVG 804
Db 721 ARLDGQFQVILVWRDLNPRSLALDPTKGYIYTWENGKPRIVAFMDGTNCMTLVKVG 780
Qy 805 RANDLTIDYADQRLYWTDLDTNMISSNMLGQERVIADLLPHFGLTQYSDIYIYWTQW 864
Db 781 RANDLTIDYADQRLYWTDLDTNMISSNMLGQERVIADLLPHFGLTQYSDIYIYWTQW 840
Qy 865 LHSTERADKTSGRNRTLIQGHLDPFVMDILVFHSSRQDGLNDCMHNNGCQGLCLAIPEGH 924
Db 841 LHSTERADKTSGRNRTLIQGHLDPFVMDILVFHSSRQDGLNDCMHNNGCQGLCLAIPEGH 900
Qy 925 RCGCASHVTLDPSSRNCSPPTTFLFSOKSAISRMIPDDOHSPLIILPLHGLRNVKAIY 984
Db 901 RCGCASHVTLDPSSRNCSPPTTFLFSOKSAISRMIPDDOHSPLIILPLHGLRNVKAIY 960
Qy 985 DPLDKFIYVVDGRQNIKRAKDDGTQPFVLTSLSQGNPDRQPHDLSIDIYSRTLFWTCEA 1044
Db 961 DPLDKFIYVVDGRQNIKRAKDDGTQPFVLTSLSQGNPDRQPHDLSIDIYSRTLFWTCEA 1020
Qy 1045 TTNINVRHLSGEAMGVVLRGDRQKPRAIIVNABRGVLYFTNMQDRAKIIERAAALDGTRE 1104
Db 1021 TTNINVRHLSGEAMGVVLRGDRQKPRAIIVNABRGVLYFTNMQDRAKIIERAAALDGTRE 1080
Qy 1105 VLFTTGLIRPVALLVNDTLGKLFWDADLKRISCDLSGANRLTLEDANTVQPLGLTILG 1164
Db 1081 VLFTTGLIRPVALLVNDTLGKLFWDADLKRISCDLSGANRLTLEDANTVQPLGLTILG 1140
Qy 1165 KHLIYWDROQMIERVEKTTGDKRTRIQRVAHLTGIHAEVVSLEEFSAHPCARDNGGC 1224
Db 1141 KHLIYWDROQMIERVEKTTGDKRTRIQRVAHLTGIHAEVVSLEEFSAHPCARDNGGC 1200
Qy 1225 SHICIAKGDTGTRCPCVHLVILONLLTCGEPPTCSPDQFACATGETDCITPGAWRCDFGP 1284
Db 1201 SHICIAKGDTGTRCPCVHLVILONLLTCGEPPTCSPDQFACATGETDCITPGAWRCDFGP 1260

QY 1285 ECDQSDSEGCPCVCSAAFFPCARGCQVRLRLRCDEADCQDRSDEADCDAICLNPQFCA 1344
Db 1261 ECDQSDSEGCPCVCSAAFFPCARGCQVRLRLRCDEADCQDRSDEADCDAICLNPQFCA 1320
QY 1345 SGQCVLIKQCCSDPDCIDGSDLMCEITKPPSDSDSPAHSSAIGPVIIGIILSLFVMGCVY 1404
Db 1321 SGQCVLIKQCCSDPDCIDGSDLMCEITKPPSDSDSPAHSSAIGPVIIGIILSLFVMGCVY 1380
QY 1405 FVCRVVCORVAGANGPPPHYVSGTPHVLNFTAPGSGQHGPPFTGACGKSMSSVSLM 1464
Db 1381 FVCRVVCORVAGANGPPPHYVSGTPHVLNFTAPGSGQHGPPFTGACGKSMSSVSLM 1440
QY 1465 GGRGVPLYDRNHVTVGASSSSSSSTKATLYPILNPPSPATDPSLYNMDIFYSSNIPAT 1524
Db 1441 GGRGVPLYDRNHVTVGASSSSSSSTKATLYPILNPPSPATDPSLYNMDIFYSSNIPAT 1500
QY 1525 ARPYRPIYIRGMAPPPTPCSTVDVCDSDYSASRWKASKYLLDLNSDSDPYPPPTPHSOYL 1584
Db 1501 VRPYRPIYIRGMAPPPTPCSTVDVCDSDYSASRWKASKYLLDLNSDSDPYPPPTPHSOYL 1560
QY 1585 SAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
Db 1561 SAEDSCPPSPATERSYFHLPPPPSPCTDSS 1591

RESULT 10

US-09-402-923A-43
; Sequence 43, Application US/09402923A
; Patent No. 655654
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; Hess, John W
; Caskey, Charles T
; Cox, Roger D
; Gerhold, David
; Hammond, Holly
; Hey, Patricia
; Kawaguchi, Yoshihiko
; Merriman, Tony R
; Metzker, Michael L
; TITLE OF INVENTION: No. 6555654el LDL-Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6555654th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; ZIP: VA 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/402,923A
; FILING DATE: 14-Feb-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB98/01102
; FILING DATE: 15-APR-1998
; APPLICATION NUMBER: US 60/043,553
; FILING DATE: 15-APR-1997
; APPLICATION NUMBER: US 60/048,740
; FILING DATE: 05-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: B.J.Sadoff
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 620-81
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4091
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1591 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 43:
US-09-402-923A-43
Query Match 98.5%; Score 8607; DB 4; Length 1591;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 25 CPAPAAASPLLLFANRRDRLVDAGGVKLESTIVVSGLEDAADAAVDFQPSKGAIVYWDVSE 84
Db 1 CPAPAAASPLLLFANRRDRLVDAGGVKLESTIVVSGLEDAADAAVDFQPSKGAIVYWDVSE 60
QY 85 EAIKQTYLNOTGAQVNVISGLVSPDGLACDWGKKLYWTDSETNRLEVANLNGTSRKV 144
Db 61 EAIKQTYLNOTGAQVNVISGLVSPDGLACDWGKKLYWTDSETNRLEVANLNGTSRKV 120
QY 145 LFWQDLDPQRAIALDPAHGYMYWTDWGETPRIERAGMDGSTRKIIIVDSDIYPNGLTIDL 204
Db 121 LFWQDLDPQRAIALDPAHGYMYWTDWGETPRIERAGMDGSTRKIIIVDSDIYPNGLTIDL 180
QY 205 EEQKLYWADAKLSPIHRANLDGSPQKVVEGSLTHPPFALTLSGDTLYWTDWQTRSIHACN 264
Db 181 EEQKLYWADAKLSPIHRANLDGSPQKVVEGSLTHPPFALTLSGDTLYWTDWQTRSIHACN 240
QY 265 KRTGGKKEILSALYSAMDIOVLQSOERQPPPHTRCEEDNGCGSHCLLSPSEPPFYTCAC 324
Db 241 KRTGGKKEILSALYSAMDIOVLQSOERQPPPHTRCEEDNGCGSHCLLSPSEPPFYTCAC 300
QY 325 TGVQLQDNGRTCKAGABEVLLARRTDLRLISLSTPDTDIVLQVDDIRHAIADYDPLE 384
Db 301 TGVQLQDNGRTCKAGABEVLLARRTDLRLISLSTPDTDIVLQVDDIRHAIADYDPLE 360
QY 385 GYVYWTDDVRAIRRAYLDGSGAQTLYNTEINDDPGIAVWVARNLYWTDGTDRIVETR 444
Db 361 GYVYWTDDVRAIRRAYLDGSGAQTLYNTEINDDPGIAVWVARNLYWTDGTDRIVETR 420
QY 445 LNGTSRKILVSEDLDEPRATLHPVMGLVWYTDGENPKIECANLDGQERRVLVNASLGM 504
Db 421 LNGTSRKILVSEDLDEPRATLHPVMGLVWYTDGENPKIECANLDGQERRVLVNASLGM 480
QY 505 PNGLALDLQEGKLYWGDAKTDKIEVINVDGTRKRTLLEDKLPHIFGFTLLGDFIYWTDMQ 564
Db 481 PNGLALDLQEGKLYWGDAKTDKIEVINVDGTRKRTLLEDKLPHIFGFTLLGDFIYWTDMQ 540
QY 565 RRSIERVHKVKSARDVIIDQLPDLMLGLKAVNNAKVNKVPNTPCADRNGGSHLCFFTPHATR 624
Db 541 RRSIERVHKVKSARDVIIDQLPDLMLGLKAVNNAKVNKVPNTPCADRNGGSHLCFFTPHATR 600
QY 625 CGCPIGLELLSDMKTCIVPEAFVFTSRAA IHRISLETNNNDVAI PLTGKVEASALDFDV 684
Db 601 CGCPIGLELLSDMKTCIVPEAFVFTSRAA IHRISLETNNNDVAI PLTGKVEASALDFDV 660
QY 685 SNNHIIYWDVSLKTIISRAFMNGSSVEHVVEFGLDYPGMAVDMWGNKLYWADTGNRIEV 744
Db 661 SNNHIIYWDVSLKTIISRAFMNGSSVEHVVEFGLDYPGMAVDMWGNKLYWADTGNRIEV 720
QY 745 ARLDGQFRQVLVWRDLNPRSLALDPKGYIYWTWEGGKPRIVRAFMNMGNTMVLVDKVG 804
Db 721 ARLDGQFRQVLVWRDLNPRSLALDPKGYIYWTWEGGKPRIVRAFMNMGNTMVLVDKVG 780
QY 805 RANDLTIDYADORLYWTDLTNMISSNMLGOERVIADDLPHFPGLTOYSDIYIYWTOWN 864
Db 781 RANDLTIDYADORLYWTDLTNMISSNMLGOERVIADDLPHFPGLTOYSDIYIYWTOWN 840
QY 865 LHSIERADKTSGRNRTLIQGHLDVFMVILVPHSSRQDGLNDCMNNNGCGQLCLAI PGGH 924
Db 841 LHSIERADKTSGRNRTLIQGHLDVFMVILVPHSSRQDGLNDCMNNNGCGQLCLAI PGGH 900
QY 925 RCGCASHYTLDPSSRNCSPTTFLFLFSQKSAISRMIDDDQHSPLDILPLHGLRNVAIDY 984
Db 901 RCGCASHYTLDPSSRNCSPTTFLFLFSQKSAISRMIDDDQHSPLDILPLHGLRNVAIDY 960

QY 985 DPLDKFIYWDGRQNIKRAKDDGTQPPVLTSLSQGNPDROPHDLSIDIYSRTLFWTCBA 1044
Db 961 DPLDKFIYWDGRQNIKRAKDDGTQPPVLTSLSQGNPDROPHDLSIDIYSRTLFWTCBA 1020
QY 1045 TINTINVRHLSGEAMGVVLRGDRDKPRAIVVNAERGVLVFTNMQDRAAKTIERAALDGTRE 1104
Db 1021 TINTINVRHLSGEAMGVVLRGDRDKPRAIVVNAERGVLVFTNMQDRAAKTIERAALDGTRE 1080
QY 1105 VLFTTGLIRPVAVLVONTLKGKLFWDADLRIKIESCDLSGANRLTLEDANIVQPLGLTILG 1164
Db 1081 VLFTTGLIRPVAVLVONTLKGKLFWDADLRIKIESCDLSGANRLTLEDANIVQPLGLTILG 1140
QY 1165 KHLVWIDRQOQMIERVEKTTGDKRTIIGRAVHLTGIIHAVEVSLIEEFSAPHCARDNGGC 1224
Db 1141 KHLVWIDRQOQMIERVEKTTGDKRTIIGRAVHLTGIIHAVEVSLIEEFSAPHCARDNGGC 1200
QY 1225 SHICIAKGDGTGPRCSPVHLVLLQNLITCGEPTCSPDQFACATGEIDCIPGAWCDGFP 1284
Db 1201 SHICIAKGDGTGPRCSPVHLVLLQNLITCGEPTCSPDQFACATGEIDCIPGAWCDGFP 1260
QY 1285 ECDQSDREGCVCSAAQPPCARGQCVLRLRCDEADCDQSDSDADDAICLPNQFRCA 1344
Db 1261 ECDQSDREGCVCSAAQPPCARGQCVLRLRCDEADCDQSDSDADDAICLPNQFRCA 1320
QY 1345 SQCVLIIKQCCDSFPDCIDGSDDELCEITKPSDDSPAHSSAIGPVIGIILSLFVMGGVY 1404
Db 1321 SQCVLIIKQCCDSFPDCIDGSDDELCEITKPSDDSPAHSSAIGPVIGIILSLFVMGGVY 1380
QY 1405 FVQRCVWCORVAGANGPPHEHYSGTTPHVPNFIAFGGSGHQGPTGACGKSMSSVSIM 1464
Db 1381 FVQRCVWCORVAGANGPPHEHYSGTTPHVPNFIAFGGSGHQGPTGACGKSMSSVSIM 1440
QY 1465 GGRGVPLVYDRNHVTGASSSSSSSTKATLYPPLNPPSPATDPSLYNMDFYSSNIPAT 1524
Db 1441 GGRGVPLVYDRNHVTGASSSSSSSTKATLYPPLNPPSPATDPSLYNMDFYSSNIPAT 1500
QY 1525 ARPYRPIYIRGMAPPTPCSTDCVSDYSASRWKASKYLLDLNDSDPYPPPTPHSOYL 1584
Db 1501 VRYRPIYIRGMAPPTPCSTDCVSDYSASRWKASKYLLDLNDSDPYPPPTPHSOYL 1560
QY 1585 SAEDSCPPSPATERSYFHLFPFPPSPCTDSS 1615
Db 1561 SAEDSCPPSPATERSYFHLFPFPPSPCTDSS 1591

RESULT 11

US-09-060-299-29
; Sequence 29, Application US/09060299
; Patent No. 6545137

GENERAL INFORMATION:
; APPLICANT: Todd, John A

; APPLICANT: Hess, John W

; APPLICANT: Caskey, Charles T

; APPLICANT: Cox, Roger D

; APPLICANT: Gerhold, David

; APPLICANT: Hammond, Holly

; APPLICANT: Hey, Patricia

; APPLICANT: Kawaguchi, Yoshihiko

; APPLICANT: Merriman, Tony R

; APPLICANT: Metzker, Michael L

; TITLE OF INVENTION: No. 6545137el Receptor
; NUMBER OF SEQUENCES: 455

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Nixon and Vanderhye

; STREET: 1100 No. 6545137th Glebe Road, Eighth Floor

; CITY: Arlington

; STATE: Virginia

; COUNTRY: US

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/060.299

; FILING DATE: 15-APR-1998

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/043,553

; FILING DATE: 15-APR-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/048,740

; FILING DATE: 05-JUN-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: B.J. Sadoff

; REGISTRATION NUMBER: 36,663

; REFERENCE/DOCKET NUMBER: 620-35

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 816-4091

; TELEFAX: (703) 816-4100

; INFORMATION FOR SEQ ID NO: 29:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1639 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

US-09-060-299-29

Query Match 98.1%; Score 8574; DB 4; Length 1639;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1585; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 26 PAAAAEPLLLFANRRDRLVDAGGVKLESTIVVSGLEDAADVDFQFSKGAIVYMTDVEE 85

Db 50 PSLHNSPFLLLFANRRDRLVDAGGVKLESTIVVSGLEDAADVDFQFSKGAIVYMTDVEE 109

QY 86 AIKQTYLNTQTAAYQNVVVISGLVSPDGLACDWGKLYTWTSETNRIEVANLNGTSRVL 145

Db 110 AIKQTYLNTQTAAYQNVVVISGLVSPDGLACDWGKLYTWTSETNRIEVANLNGTSRVL 169

QY 146 FWQDLQDPRALDPAHGYMYTWDGTPRIERAGMGSTRKIIIVDSIYPNGLTIDLE 205

Db 170 FWQDLQDPRALDPAHGYMYTWDGTPRIERAGMGSTRKIIIVDSIYPNGLTIDLE 229

QY 206 EQKLYWADAKLSFTHRANLDGSPQKVVVEGSLTHPFFALTLSGDTLYTWTDTWTRIHCNK 265

Db 230 EQKLYWADAKLSFTHRANLDGSPQKVVVEGSLTHPFFALTLSGDTLYTWTDTWTRIHCNK 289

QY 266 RTGGRKEILLSALYSPMDIQVLSQERQPPFHTRCEDNGGCSHLCLLSPSPFFYTCACT 325

Db 290 RTGGRKEILLSALYSPMDIQVLSQERQPPFHTRCEDNGGCSHLCLLSPSPFFYTCACT 349

QY 326 GVQLQDNGRTCKAGAEVLLIARLTDLRRISLDPDFTDVLQVDDIRHATAIDYDPLEG 385

Db 350 GVQLQDNGRTCKAGAEVLLIARLTDLRRISLDPDFTDVLQVDDIRHATAIDYDPLEG 409

QY 386 VYVWTDDEVRAIRRAYLDGSGAOTLVNTEINDPGIADVWVARLYTWTGTDTRIEVTREL 445

Db 410 VYVWTDDEVRAIRRAYLDGSGAOTLVNTEINDPGIADVWVARLYTWTGTDTRIEVTREL 469

QY 446 NGTSRKILVSEDLDDEPPRAIALHPVMGLMYTWDNGENPKIECANLDQERRVNVNASLWGP 505

Db 470 NGTSRKILVSEDLDDEPPRAIALHPVMGLMYTWDNGENPKIECANLDQERRVNVNASLWGP 529

QY 506 NGLALDIQEGKLYWGDADTKIEVINVDGTPKRTLLDKLPHIFGFTLLGDFIYWTWQOR 565

Db 530 NGLALDIQEGKLYWGDADTKIEVINVDGTPKRTLLDKLPHIFGFTLLGDFIYWTWQOR 589

QY 566 RSIERVHKVKASRDVVIIDQLPDLMLGKAVNVAKVVGTPNCPADRNGGCSHLCLFFTPHATRC 625

Db 590 RSIERVHKVKASRDVVIIDQLPDLMLGKAVNVAKVVGTPNCPADRNGGCSHLCLFFTPHATRC 649

QY 626 GCPIGLELLSDMKTICIVPEAFIVFTSPAAIHRISLENNNDVAIPLTGVKEASALDSDVS 685

Db 650 GCPIGLELLSDMKTICIVPEAFIVFTSPAAIHRISLENNNDVAIPLTGVKEASALDSDVS 709

Qy 686 NNHIYWTVDVSLKTISSAFMGSSVEHVVEGLDYPGNAVDWVGKLYWADTGNRIEVA 745
Db 710 NNHIYWTVDVSLKTISSAFMGSSVEHVVEGLDYPGNAVDWVGKLYWADTGNRIEVA 769
Qy 746 RLDGQFROVLVWRDLNPRSLALDPTKGYIYWTWGGKPRIVRAFMDGTNCWTLVDKVR 805
Db 770 RLDGQFROVLVWRDLNPRSLALDPTKGYIYWTWGGKPRIVRAFMDGTNCWTLVDKVR 829
Qy 806 ANDLTIDYADQRLVWDLTWNMISSNMLGQERVVIADLPHPGTGOYSYIYWTWNL 865
Db 830 ANDLTIDYADQRLVWDLTWNMISSNMLGQERVVIADLPHPGTGOYSYIYWTWNL 889
Qy 866 HSIERADTSGNRNRTLIQHLDFVMDILVHFSSRODGLNDCHWNNQCGQLCLAIPEGHR 925
Db 890 HSIERADTSGNRNRTLIQHLDFVMDILVHFSSRODGLNDCHWNNQCGQLCLAIPEGHR 949
Qy 926 CGCASHYTLDPSSRNCSPPTTFLPFSQSAISRMIPDQHSFDPDLPLHLGLRNKVAIDYD 985
Db 950 CGCASHYTLDPSSRNCSPPTTFLPFSQSAISRMIPDQHSFDPDLPLHLGLRNKVAIDYD 1009
Qy 986 PLDKFIYWDGRONIKRAKDGTQPFVLTSLSQGNPDQPHDLSIDIYSRTLFWTCEAT 1045
Db 1010 PLDKFIYWDGRONIKRAKDGTQPFVLTSLSQGNPDQPHDLSIDIYSRTLFWTCEAT 1069
Qy 1046 NTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERGVLFTNNQDRAAKIERAALDGTREV 1105
Db 1070 NTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERGVLFTNNQDRAAKIERAALDGTREV 1129
Qy 1106 LFTTGLIRPVAVDNTLGLKFWVDADLKRIESCDSLGNRLTLEDANIVQPLGLTLGK 1165
Db 1130 LFTTGLIRPVAVDNTLGLKFWVDADLKRIESCDSLGNRLTLEDANIVQPLGLTLGK 1189
Qy 1166 HLYMIDROQMIERVEKTTGDKRTIQRVAVHLTGIIHAEVSVLEEFSAHPCARDNGCS 1225
Db 1190 HLYMIDROQMIERVEKTTGDKRTIQRVAVHLTGIIHAEVSVLEEFSAHPCARDNGCS 1249
Qy 1226 HICIAKGDTGTRCPCPVHLVLLQNLITCGEPPTCPDQFACATGEIDICPGAWRCDFPE 1285
Db 1250 HICIAKGDTGTRCPCPVHLVLLQNLITCGEPPTCPDQFACATGEIDICPGAWRCDFPE 1309
Qy 1286 CDDQSDGEGCVCSAAPPCARGCQVLRRCDEADCCQDRSDEADCDALCLPNQFRCS 1345
Db 1310 CDDQSDGEGCVCSAAPPCARGCQVLRRCDEADCCQDRSDEADCDALCLPNQFRCS 1369
Qy 1346 GQCVLIKQCCDFDCIDGSDLMCEITKPPSDSPAHSSAIGVIGIILSLFWMGVYF 1405
Db 1370 GQCVLIKQCCDFDCIDGSDLMCEITKPPSDSPAHSSAIGVIGIILSLFWMGVYF 1429
Qy 1406 VCQVVCQRYAGANGPPFHEVYVSGTPHVPFLNFIAPGSGHQHGFPGIACGKSMGSSVSLMG 1465
Db 1430 VCQVVCQRYAGANGPPFHEVYVSGTPHVPFLNFIAPGSGHQHGFPGIACGKSMGSSVSLMG 1489
Qy 1466 GRGGVPLYDRNHVTGASSSSSSSTKATLYPILNPPSPATDPSLYNMDMFPYSNIPATA 1525
Db 1490 GRGGVPLYDRNHVTGASSSSSSSTKATLYPILNPPSPATDPSLYNMDMFPYSNIPATA 1549
Qy 1526 RPYRPIIRGHAPPTTCCSTDCVSDYSASRWKASKYIYDLNLSDDPYPPPTPHSQYLS 1585
Db 1550 RPYRPIIRGHAPPTTCCSTDCVSDYSASRWKASKYIYDLNLSDDPYPPPTPHSQYLS 1609
Qy 1586 AEDSCPPSPATERSYFHLFPPTTSPCTDSS 1615
Db 1610 AEDSCPPSPATERSYFHLFPPTTSPCTDSS 1639

RESULT 12

US-09-402-923A-29
; Sequence 29, Application US/09402923A
; Patent No. 6555654
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; Hess, John W
; Caskey, Charles T

Cox, Roger D
Gerhold, David
Hammond, Holly
Hey, Patricia
Kawaguchi, Yoshihiko
Merriman, Tony R
Metzker, Michael L
TITLE OF INVENTION: No. 6555654e1 LDL-Receptor
NUMBER OF SEQUENCES: 455
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon and Vanderhye
STREET: 1100 No. 6555654th Glebe Road, Eighth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: US
ZIP: VA 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402.923A
FILING DATE: 14-Feb-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB98/01102
FILING DATE: 15-APR-1998
APPLICATION NUMBER: US 60/043,553
FILING DATE: 15-APR-1997
APPLICATION NUMBER: US 60/048,740
FILING DATE: 05-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: B.J.Sadoff
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 620-81
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4091
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 1639 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-09-402-923A-29

Query Match 98.1%; Score 8574; DB 4; Length 1639;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1585; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 26 PAPAAASPLLLPANRRDRLVYDAGVVKLESTIVVSGLEDAADVPQPSKGAVYWTVDSEE 85
Db 50 PSLHNSPPLLLPANRRDRLVYDAGVVKLESTIVVSGLEDAADVPQPSKGAVYWTVDSEE 109
Qy 86 AIKOTYLNQTAAGVQNVVISGLVSPDGLACDWGKLYWTDSETNRIEVLNLTGTSRKVL 145
Db 110 AIKOTYLNQTAAGVQNVVISGLVSPDGLACDWGKLYWTDSETNRIEVLNLTGTSRKVL 169
Qy 146 FWQDLQDPAIALDPAHGYMYWTWGTETPRIERAGMDGSTRKIIIVDSDIYWPNGLTIDLE 205
Db 170 FWQDLQDPAIALDPAHGYMYWTWGTETPRIERAGMDGSTRKIIIVDSDIYWPNGLTIDLE 229
Qy 206 EOKLYWADAKLSFIHRANLDGSPQKVVEGSLTHPPFALTLSGDTLYWTDWQTRSHACNK 265
Db 230 EOKLYWADAKLSFIHRANLDGSPQKVVEGSLTHPPFALTLSGDTLYWTDWQTRSHACNK 289
Qy 266 RTGKRKEILSALYSPMDIOVLISEROPFFHTRCEEENGCGSHLCLLSPSPFFYTCAPT 325
Db 290 RTGKRKEILSALYSPMDIOVLISEROPFFHTRCEEENGCGSHLCLLSPSPFFYTCAPT 349
Qy 326 GVQLQDNGRGTCKAGAEVLLARRTDLRRISLDTPDFTDIVLQVDDIRHAIAIDYDPLEG 385
Db 350 GVQLQDNGRGTCKAGAEVLLARRTDLRRISLDTPDFTDIVLQVDDIRHAIAIDYDPLEG 409

QY 386 YVYWTDDVRAIRRAYLDGSGAQLVNTIINDPGIAVDWVARNLWTGTGTRIEVTRL 445
Db 410 YVYWTDDVRAIRRAYLDGSGAQLVNTIINDPGIAVDWVARNLWTGTGTRIEVTRL 469
QY 446 NGTSRKILVSDLEBPRAIALHPVWGLMYTDWGENPKIECANLDGQERRVLVNASLGWP 505
Db 470 NGTSRKILVSDLEBPRAIALHPVWGLMYTDWGENPKIECANLDGQERRVLVNASLGWP 529
QY 506 NGLALDLQEGKLYGDAKTIKIEVINVDGTRKRLLEDKPHIFGFTLLGDFIYWTWQOR 565
Db 530 NGLALDLQEGKLYGDAKTIKIEVINVDGTRKRLLEDKPHIFGFTLLGDFIYWTWQOR 589
QY 566 RSIERVHKVKSARDVIIDQLPDLMLGLKAVNVAKVGTNPNPCADNRNGGCSHLCTFFTPHATRC 625
Db 590 RSIERVHKVKSARDVIIDQLPDLMLGLKAVNVAKVGTNPNPCADNRNGGCSHLCTFFTPHATRC 649
QY 626 GCPGLELLSDMKTCIVPEAFVFTSRAAIIHRSLETNNNDVAIPLTGCVKEASALDFDVS 685
Db 650 GCPGLELLSDMKTCIVPEAFVFTSRAAIIHRSLETNNNDVAIPLTGCVKEASALDFDVS 709
QY 686 NNHIYWTDVSLKTSRAFWNGSSVEHVVEFGLDYPEGMAVDMGKNLYWADTGNRIEVA 745
Db 710 NNHIYWTDVSLKTSRAFWNGSSVEHVVEFGLDYPEGMAVDMGKNLYWADTGNRIEVA 769
QY 746 RLDGQFRQVLVWRDLNDRSLALDPTKGYIYWTWGGKPRIVRAFPMGTNCTMLVDKVR 805
Db 770 RLDGQFRQVLVWRDLNDRSLALDPTKGYIYWTWGGKPRIVRAFPMGTNCTMLVDKVR 829
QY 806 ANDLTIDYADQRLYWTDLTNMIESNNMLGQERVVIADLPHFGLTGOYSYIYWTWNL 865
Db 830 ANDLTIDYADQRLYWTDLTNMIESNNMLGQERVVIADLPHFGLTGOYSYIYWTWNL 889
QY 866 HSIERADTSGNRNLTIOGHLDVMDILVHSSRDGLNDCHMNNQCQGLCLAIPEGHR 925
Db 890 HSIERADTSGNRNLTIOGHLDVMDILVHSSRDGLNDCHMNNQCQGLCLAIPEGHR 949
QY 926 CCASHYLDPSRNCSPPTFLFSQSAISRMIPDQHSPLILPLHGLRNKVAIDYD 985
Db 950 CCASHYLDPSRNCSPPTFLFSQSAISRMIPDQHSPLILPLHGLRNKVAIDYD 1009
QY 986 PLDKFIYWDGQNKIRAKDGTQFVLTSLSQGNPDRQPHDLSDIYSRTLFWTCBAT 1045
Db 1010 PLDKFIYWDGQNKIRAKDGTQFVLTSLSQGNPDRQPHDLSDIYSRTLFWTCBAT 1069
QY 1046 NTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERGVLVYFTNMQDRAAKIERAALDGTREY 1105
Db 1070 NTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERGVLVYFTNMQDRAAKIERAALDGTREY 1129
QY 1106 LFTTGLIRFVALVNDTLGKLFWDADLKRIBSCDLSGANRLTLTDANIVQPLGTLILGK 1165
Db 1130 LFTTGLIRFVALVNDTLGKLFWDADLKRIBSCDLSGANRLTLTDANIVQPLGTLILGK 1189
QY 1166 HLYWIDRQOQMLERVEKTTGDKRTRIQGRVAHLTGTHAVEEVSLEEFSAHPCARDNGGCS 1225
Db 1190 HLYWIDRQOQMLERVEKTTGDKRTRIQGRVAHLTGTHAVEEVSLEEFSAHPCARDNGGCS 1249
QY 1226 HICIAKGDGTPRCSPVHLVLLQNLITCGEPTCSPDQACATGIDCIPGAWRCGPRPE 1285
Db 1250 HICIAKGDGTPRCSPVHLVLLQNLITCGEPTCSPDQACATGIDCIPGAWRCGPRPE 1309
QY 1286 CDDQSEEGCPVCSAAQFFCARGQCVDLRLRCGDGADCCDRSDEADCDALCLPNQFRCAS 1345
Db 1310 CDDQSEEGCPVCSAAQFFCARGQCVDLRLRCGDGADCCDRSDEADCDALCLPNQFRCAS 1369
QY 1346 GQCVLIIKQCDSPDPCIDGSDMLCEITKPPSDSPHSAISGIPVIGIILSLFWMGVYF 1405
Db 1370 GQCVLIIKQCDSPDPCIDGSDMLCEITKPPSDSPHSAISGIPVIGIILSLFWMGVYF 1429
QY 1406 VQORVVCQRYAGANGFPFHYEYSGTPHPLNFTAPGSGOHPPTGTACGKSNMSSVSLMG 1465
Db 1430 VQORVVCQRYAGANGFPFHYEYSGTPHPLNFTAPGSGOHPPTGTACGKSNMSSVSLMG 1489

QY 1466 GRGCVPLYDRNHVTGASSSSSSSTKATLYPPILNPPSPATDPSLYNMDMFYSSNIPATA 1525
Db 1490 GRGCVPLYDRNHVTGASSSSSSSTKATLYPPILNPPSPATDPSLYNMDMFYSSNIPATA 1549
QY 1526 RPVRPYIIRGMAPPTTPCSTDVCDSDYSASRWKASKYIYLDLNSDSDPYPPPTPHSQYLS 1585
Db 1550 RPVRPYIIRGMAPPTTPCSTDVCDSDYSASRWKASKYIYLDLNSDSDPYPPPTPHSQYLS 1609
QY 1586 AEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
Db 1610 AEDSCPPSPATERSYFHLPPPPSPCTDSS 1639

RESULT 13

US-09-060-299-39
; Sequence 39, Application US/09060299
; Patent No. 6545137
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; APPLICANT: Hesse, John W
; APPLICANT: Caskey, Charles T
; APPLICANT: Cox, Roger D
; APPLICANT: Gerhold, David
; APPLICANT: Hammond, Holly
; APPLICANT: Hey, Patricia
; APPLICANT: Kawaguchi, Yoshihiko
; APPLICANT: Merriman, Tony R
; APPLICANT: Metzker, Michael L
; TITLE OF INVENTION: No. 6545137el Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6545137th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,299
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/043,553
; FILING DATE: 15-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,740
; FILING DATE: 05-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: B.J.Sadoff
; REGISTRATION NUMBER: 36,663
; REFERENCE/POCKET NUMBER: 620-35
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4091
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1584 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-060-299-39

Query Match 98.0%; Score 8568; DB 4; Length 1584;

Best Local Similarity 99.9%; Fred. No. 0; Indels 0; Gaps 0;
Matches 1583; Conservative 0; Mismatches 1;

QY 32 SPLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQFSKGAVYWTVDVSEEAIKQTY 91
Db 1 SPLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQFSKGAVYWTVDVSEEAIKQTY 60

QY 92 LNQTGAAYONVYVIGSLVSPDGLACDWGKLYWTDSETNRIEVANLNGTSRKVLFWQDLD 151
DB 61 LNQTGAAYONVYVIGSLVSPDGLACDWGKLYWTDSETNRIEVANLNGTSRKVLFWQDLD 120
QY 152 OPRAIALDPANGYMYWTDWGETPRIERAGMGSTRKIIIVDSDIYWPNGLTIDLEBQKLYW 211
DB 121 QPRAIALDPANGYMYWTDWGETPRIERAGMGSTRKIIIVDSDIYWPNGLTIDLEBQKLYW 180
QY 212 ADAXLSFTHIRANLDGSPQKVVVEGSLTHPPFALTJSGDLYWTDWOTRSHACNKRKTGGR 271
DB 181 ADAXLSFTHIRANLDGSPQKVVVEGSLTHPPFALTJSGDLYWTDWOTRSHACNKRKTGGR 240
QY 272 KEILSALYSPMDIQVLSQEROPFFHTRCEEDNGGCSHLCLLSPSEFFYTACPTGVQLQD 331
DB 241 KEILSALYSPMDIQVLSQEROPFFHTRCEEDNGGCSHLCLLSPSEFFYTACPTGVQLQD 300
QY 332 NGRTCKAGAEVLLARLTDLRRISLDPDFTDVLQVDDIRHAIAIDYDPLEGVYWTD 391
DB 301 NGRTCKAGAEVLLARLTDLRRISLDPDFTDVLQVDDIRHAIAIDYDPLEGVYWTD 360
QY 392 DEVRAIRRAYLDGSGAQTIVNTEINDPGIAVDWVARNLYWTDGTDRIEVTRLNGTSRK 451
DB 361 DEVRAIRRAYLDGSGAQTIVNTEINDPGIAVDWVARNLYWTDGTDRIEVTRLNGTSRK 420
QY 452 ILVSEDLDEPRAIALHPVMGLMYWTDWGENPKIECANLDGQERRVYNASLWPNGLALD 511
DB 421 ILVSEDLDEPRAIALHPVMGLMYWTDWGENPKIECANLDGQERRVYNASLWPNGLALD 480
QY 512 LOEGKLYWGDAKTIEVINVDGTRKRTLLBCKLPHIFPGFTLLGDFIYWTWQRRSIRV 571
DB 481 LOEGKLYWGDAKTIEVINVDGTRKRTLLBCKLPHIFPGFTLLGDFIYWTWQRRSIRV 540
QY 572 HKVKSARDVIIDQLPDLMLGKAVNAVGVNCPADRNGGCSHLCTFFTPHATRCGCPGL 631
DB 541 HKVKSARDVIIDQLPDLMLGKAVNAVGVNCPADRNGGCSHLCTFFTPHATRCGCPGL 600
QY 632 ELLSDMKTCIIVPEAFVFTSRAAIIHRISLETNNNDVAIPLTGVKEASALDFDVSNHHIY 691
DB 601 ELLSDMKTCIIVPEAFVFTSRAAIIHRISLETNNNDVAIPLTGVKEASALDFDVSNHHIY 660
QY 692 TDVSLKTIISRAFPMGSSVEHVVEFGLDYPEGMAVDWGMKNLYWADTGTNRIEVARLDGQF 751
DB 661 TDVSLKTIISRAFPMGSSVEHVVEFGLDYPEGMAVDWGMKNLYWADTGTNRIEVARLDGQF 720
QY 752 RQVLVWRDLNPRISALDPTKGYIYWTWEGCKPRIVRAFMDDGTNCMTLVDKVGANDLTI 811
DB 721 RQVLVWRDLNPRISALDPTKGYIYWTWEGCKPRIVRAFMDDGTNCMTLVDKVGANDLTI 780
QY 812 DYADORLYWTDLDTNMISSNMLGOERVVIADDLPHFPGLTQYSDIYIYWTWDLNLSIERA 871
DB 781 DYADORLYWTDLDTNMISSNMLGOERVVIADDLPHFPGLTQYSDIYIYWTWDLNLSIERA 840
QY 872 DKTSGRNRTLQGLHDFWMDILVPHSSRQDGLNOMHNGCGQLCLAIPOGHRGCGASH 931
DB 841 DKTSGRNRTLQGLHDFWMDILVPHSSRQDGLNOMHNGCGQLCLAIPOGHRGCGASH 900
QY 932 YTLDPSSRNCSPPTTFLFSSOKSAISRMIPDOHSPDLILPLHGLRNKVAIDYDPLDKFI 991
DB 901 YTLDPSSRNCSPPTTFLFSSOKSAISRMIPDOHSPDLILPLHGLRNKVAIDYDPLDKFI 960
QY 992 YWVDGRQNIKRAKDDGTGPPVLTLSQGNPDROPHDLSIDIYSRTLFWTCBATNTINVH 1051
DB 961 YWVDGRQNIKRAKDDGTGPPVLTLSQGNPDROPHDLSIDIYSRTLFWTCBATNTINVH 1020
QY 1052 RLSGEAMGVVLGRDRDKPRAIIVNAERGILYFTNMQRRAAKIERAALDGTREVERLFTTGL 1111
DB 1021 RLSGEAMGVVLGRDRDKPRAIIVNAERGILYFTNMQRRAAKIERAALDGTREVERLFTTGL 1080
QY 1112 IRPVALVVDNTLGLKFWVDADLKEIESCDLSGANRLTLEDANIYQPLGLTILGKHLWID 1171
DB 1081 IRPVALVVDNTLGLKFWVDADLKEIESCDLSGANRLTLEDANIYQPLGLTILGKHLWID 1140
QY 1172 RQQMIEVERKTTGDKRTRIQRVAHLTGIHAEVSELSAHPCARDNGGCSHICIAK 1231

DB 1141 RQQMIEVERKTTGDKRTRIQRVAHLTGIHAEVSELSAHPCARDNGGCSHICIAK 1200
QY 1232 GDGTPRCSCPVLVLLQNLITCGEPPTCSPOQFACATGEIDICIPGAWRCDFPPECDDQSD 1291
DB 1201 GDGTPRCSCPVLVLLQNLITCGEPPTCSPOQFACATGEIDICIPGAWRCDFPPECDDQSD 1260
QY 1292 BEGPCVCSAAOPPCARGOCVDLRLRCDEADQDRSDDEADDAICLPNOFRCSAQCVLI 1351
DB 1261 BEGPCVCSAAOPPCARGOCVDLRLRCDEADQDRSDDEADDAICLPNOFRCSAQCVLI 1320
QY 1352 KQCCDSPPDCIDGSDDELMEITKPPSDSPAHSSAIGVIGIILSLFVMGGYFVCQVW 1411
DB 1321 KQCCDSPPDCIDGSDDELMEITKPPSDSPAHSSAIGVIGIILSLFVMGGYFVCQVW 1380
QY 1412 CORVAGANGPPHEVYSGTTPHVLNFIAPGSGORHPPFTGIACGKSMSSVSLMGGRGV 1471
DB 1381 CORVAGANGPPHEVYSGTTPHVLNFIAPGSGORHPPFTGIACGKSMSSVSLMGGRGV 1440
QY 1472 LYDRNHVTGASSSSSTKATLYPILNPPSPATDPSLYNMDMYSSNIPATAPYRPY 1531
DB 1441 LYDRNHVTGASSSSSTKATLYPILNPPSPATDPSLYNMDMYSSNIPATAPYRPY 1500
QY 1532 IIRGWAPPTTCTSDVCDSDYSASRWKASKYLLDNLSDSDPYPPPTPHSQYLSAEDSCP 1591
DB 1501 IIRGWAPPTTCTSDVCDSDYSASRWKASKYLLDNLSDSDPYPPPTPHSQYLSAEDSCP 1560
QY 1592 PSPATERSYFHLFPPPPSPCTDSS 1615
DB 1561 PSPATERSYFHLFPPPPSPCTDSS 1584

RESULT 14

US-09-402-923A-39
; Sequence 39, Application US/09402923A
; Patent No. 6555654
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; Hess, John W
; Caskey, Charles T
; Cox, Roger D
; Gerhold, David
; Hammond, Holly
; Hey, Patricia
; Kawaguchi, Yoshihiko
; Merriman, Tony R
; Metzker, Michael L
; TITLE OF INVENTION: No. 6555654e1 LDL-Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6555654th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; ZIP: VA 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/402,923A
; FILING DATE: 14-Feb-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB98/01102
; FILING DATE: 15-APR-1998
; APPLICATION NUMBER: US 60/043,553
; FILING DATE: 15-APR-1997
; APPLICATION NUMBER: US 60/048,740
; FILING DATE: 05-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: B.J.Sadoff

REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 620-81
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4091
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 1584 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-09-402-923A-39

Query Match 98.0%; Score 8568; DB 4; Length 1584;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 1583; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 32 SPLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQSKGAVYWDVSEEAIKQTY 91
DB 1 SPLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQSKGAVYWDVSEEAIKQTY 60

QY 92 LNQTGAQVNVVISGLVSPDGLACDWGKLYWTDSETNRIEVANLNGTSRKVLFWQDLD 151
DB 61 LNQTGAQVNVVISGLVSPDGLACDWGKLYWTDSETNRIEVANLNGTSRKVLFWQDLD 120

QY 152 QPRAIALDPAHGYMYTWDGETPRIERAGMDGSTRKIIIVDSIIYWPNGLTIDLBEOKLYW 211
DB 121 QPRAIALDPAHGYMYTWDGETPRIERAGMDGSTRKIIIVDSIIYWPNGLTIDLBEOKLYW 180

QY 212 ADAKLSFTHRANLDGSKVQVVEGSLTHPEALTLSGDTLYWTDWTRSIHACNKRKTGGR 271
DB 181 ADAKLSFTHRANLDGSKVQVVEGSLTHPEALTLSGDTLYWTDWTRSIHACNKRKTGGR 240

QY 272 KEILSALYSPMDIQVLSQERQPFHTRCEEDNGGCSHLCLLSPSPFFTCACPTGVQLQD 331
DB 241 KEILSALYSPMDIQVLSQERQPFHTRCEEDNGGCSHLCLLSPSPFFTCACPTGVQLQD 300

QY 332 NGRTCKAGAEVLLIARRTDLRIISLDPDFDVIQVDDIRHATAIDYDPLEGVVYWD 391
DB 301 NGRTCKAGAEVLLIARRTDLRIISLDPDFDVIQVDDIRHATAIDYDPLEGVVYWD 360

QY 392 DEVRAIRAYLDGSAQTLVNTIENDPGIAVDVARNLYWTDGTGRIEVRNLNGTSRK 451
DB 361 DEVRAIRAYLDGSAQTLVNTIENDPGIAVDVARNLYWTDGTGRIEVRNLNGTSRK 420

QY 452 ILVSEDLDEPRAIALHPVMGLMYTWDGPNKIECANLDGQERRVLVNASLGWPNGLALD 511
DB 421 ILVSEDLDEPRAIALHPVMGLMYTWDGPNKIECANLDGQERRVLVNASLGWPNGLALD 480

QY 512 LOEGKLYWGDAKTDKLEIVNVDTGKRTLLDKLPHITGFTLLGDFIYWDWQRSIERV 571
DB 481 LOEGKLYWGDAKTDKLEIVNVDTGKRTLLDKLPHITGFTLLGDFIYWDWQRSIERV 540

QY 572 HKVKASRDVIIIDQLPDLMLGLKAVNAKVVGTVNTPCADRNGGCSHLCTFPHATRCGPIGL 631
DB 541 HKVKASRDVIIIDQLPDLMLGLKAVNAKVVGTVNTPCADRNGGCSHLCTFPHATRCGPIGL 600

QY 632 ELLSDMKTCIVPEAFIVFTSRAAIHRIISLETNNNDVAIPLTGKESALDFDVSNNHIYW 691
DB 601 ELLSDMKTCIVPEAFIVFTSRAAIHRIISLETNNNDVAIPLTGKESALDFDVSNNHIYW 660

QY 692 TDVSLKTIISRAFNGSSVEHVEFGLDYPEGMAVDMWGNKLYWADTGTNRIEVARLDGOF 751
DB 661 TDVSLKTIISRAFNGSSVEHVEFGLDYPEGMAVDMWGNKLYWADTGTNRIEVARLDGOF 720

QY 752 RQVLVWRDLNPRSLALDPTKGIYIYTWEGGKPRIIVRAFMDGTNCMTLVKVGRLNDLTI 811
DB 721 RQVLVWRDLNPRSLALDPTKGIYIYTWEGGKPRIIVRAFMDGTNCMTLVKVGRLNDLTI 780

QY 812 DYADQRLYWTDLDTNNIESNMLQORVVIADLPHPGITQYSYDIYWDNWLHSIERA 871
DB 781 DYADQRLYWTDLDTNNIESNMLQORVVIADLPHPGITQYSYDIYWDNWLHSIERA 840

QY 872 DKTSGRNRTLIQGHLDPMVDILVFHSSRODGLNDCHMNNQCCGOLCLAIPEGHRCGCASH 931
DB 841 DKTSGRNRTLIQGHLDPMVDILVFHSSRODGLNDCHMNNQCCGOLCLAIPEGHRCGCASH 900

QY 932 YTLDPSSRNCSPPTTFLFSQKSAISRMIPDDQHSPLILPLHGLRNVAIDAIDYDPLDKFI 991
DB 901 YTLDPSSRNCSPPTTFLFSQKSAISRMIPDDQHSPLILPLHGLRNVAIDAIDYDPLDKFI 960

QY 992 YWVDGRONIKRAKDDGTQPFVLTSLSQGNPDQPHDLSIDIYSRTLFWTCEATNTINHV 1051
DB 961 YWVDGRONIKRAKDDGTQPFVLTSLSQGNPDQPHDLSIDIYSRTLFWTCEATNTINHV 1020

QY 1052 RLSGEAMGVVLGRDRKPRAIVVNAERGILYFTNMODRAAKIERAAALDGTREVLFTTGL 1111
DB 1021 RLSGEAMGVVLGRDRKPRAIVVNAERGILYFTNMODRAAKIERAAALDGTREVLFTTGL 1080

QY 1112 IRPVALLVVDNTLGLFWVDADLKRIESCDLSGANRLTLEDANIVQPLGLTILGKHLWID 1171
DB 1081 IRPVALLVVDNTLGLFWVDADLKRIESCDLSGANRLTLEDANIVQPLGLTILGKHLWID 1140

QY 1172 RQOQMIERVEKTTGDKETRIQGRVAHLTGTHAVEEVSLSEEFSAHPCARDNGGCSHICIAK 1231
DB 1141 RQOQMIERVEKTTGDKETRIQGRVAHLTGTHAVEEVSLSEEFSAHPCARDNGGCSHICIAK 1200

QY 1232 GDGTPRCSVHLVLLQNLTLTCCGEPPTCSPPQFACATGEIDCIPGAWRCDFPCCDQSD 1291
DB 1201 GDGTPRCSVHLVLLQNLTLTCCGEPPTCSPPQFACATGEIDCIPGAWRCDFPCCDQSD 1260

QY 1292 BEGCPVCSAAQFPCCARGQCVDLRLRCDEADQDRSDEADCAICLPNQFRCASQCQVLI 1351
DB 1261 BEGCPVCSAAQFPCCARGQCVDLRLRCDEADQDRSDEADCAICLPNQFRCASQCQVLI 1320

QY 1352 KQCCDSPPDCIDGDELWCETTKPPSDSPHSSAIGPVGIIISLFWMGVYVFCQRVV 1411
DB 1321 KQCCDSPPDCIDGDELWCETTKPPSDSPHSSAIGPVGIIISLFWMGVYVFCQRVV 1380

QY 1412 CORVAGANGPPPEHYVSGTHVPLNFIAPGSGHPTGTIACGSKMSSVSLMGGRGVVP 1471
DB 1381 CORVAGANGPPPEHYVSGTHVPLNFIAPGSGHPTGTIACGSKMSSVSLMGGRGVVP 1440

QY 1472 LYDRNHVTGASSSSSSSTKATLYPILNPPSPATDPSLYNMDMFYSSNIPATVPYRPY 1531
DB 1441 LYDRNHVTGASSSSSSSTKATLYPILNPPSPATDPSLYNMDMFYSSNIPATVPYRPY 1500

QY 1532 IIRGMWAPPTTCTDVCDSYASRWKASKYKYLIDNSDSDPYPPPTPHSYLSAEDSCP 1591
DB 1501 IIRGMWAPPTTCTDVCDSYASRWKASKYKYLIDNSDSDPYPPPTPHSYLSAEDSCP 1560

QY 1592 PSPATERSYFHLFPPPPSPCTDSS 1615
DB 1561 PSPATERSYFHLFPPPPSPCTDSS 1584

RESULT 15
US-09-060-299-42
; Sequence 42, Application US/09060299
; Patent No. 6545137
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; APPLICANT: Hess, John W
; APPLICANT: Caskey, Charles T
; APPLICANT: Cox, Roger D
; APPLICANT: Gerhold, David
; APPLICANT: Hammond, Holly
; APPLICANT: Hey, Patricia
; APPLICANT: Kawaguchi, Yoshihiko
; APPLICANT: Merriman, Tony R
; APPLICANT: Metzker, Michael L
; TITLE OF INVENTION: No. 6545137el Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6545137th Glebe Road, Eighth Floor

CITY: Arlington
STATE: Virginia
COUNTRY: US
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/060,299
FILING DATE: 15-APR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/043,553
FILING DATE: 15-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/048,740
FILING DATE: 05-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: B.J.Sadoff
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 620-35
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4091
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 1614 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-060-299-42

Query Match 94.5%; Score 8259.5; DB 4; Length 1614;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1519; Conservative 41; Mismatches 51; Indels 7; Gaps 2;

Qy 1 MEAAP---PGPPWPLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGGVKLESTI 57
Db 1 METAPTRAPPPPPPLLLLLLVLSL---VPAASPLLLFANRRDRLVLDAGGVKLESTI 56

Qy 58 VVSGLEDAAAVDFQSGKAVTWDVSEBAIKQTYLNQTGAQVQNVVLSGLVSPDGLACDW 117
Db 57 VASGLEDAAAVDFQSGKAVTWDVSEBAIKQTYLNQTGAQVQNVVLSGLVSPDGLACDW 116

Qy 118 VGKLYTWDSTNRLEVANLNGTSRKVLFWODLQDPAIALDPAHGYMYTWDGSTRPTE 177
Db 117 VGKLYTWDSTNRLEVANLNGTSRKVLFWODLQDPAIALDPAHGYMYTWDGSTRPTE 176

Qy 178 RAGMDGSTRKIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPFRQKVEGSL 237
Db 177 RAGMDGSTRKIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPFRQKVEGSL 236

Qy 238 THPPALTLSGDTLVTWQWTSIIHACNKTGCKEILSALYSMDIQVLQSERQPPFHT 297
Db 237 THPPALTLSGDTLVTWQWTSIIHACNKTGCKEILSALYSMDIQVLQSERQPPFHT 296

Qy 298 RCEBDNGCSHLCLLSPEPYTCACPTGVQLOQNGRTCKAGAEVLLIARTRDLRLSL 357
Db 297 RCEBDNGCSHLCLLSPEPYTCACPTGVQLOQNGRTCKAGAEVLLIARTRDLRLSL 356

Qy 358 DTPDFTDVLQVDDIRHAIAIDYDPLEGYVYVTDDEVRAIRRAYLDGSGAQLVNTIEND 417
Db 357 DTPDFTDVLQVDDIRHAIAIDYDPLEGYVYVTDDEVRAIRRAYLDGSGAQLVNTIEND 416

Qy 418 PDGIADVWAAANLVTGTDRIEIVRLNGTSRKILVSEDLDEPRALHPVGMGLMYTWD 477
Db 417 PDGIADVWAAANLVTGTDRIEIVRLNGTSRKILVSEDLDEPRALHPVGMGLMYTWD 476

Qy 478 WGENPKIECANLDCQERRVLVNASIGWPNGLALDLQEGKLYWDAKTDKIEVINVDGTR 537
Db 477 WGENPKIECANLDCQERRVLVNTSLGWPNGLALDLQEGKLYWDAKTDKIEVINVDGTR 536

Qy 538 RTLLLEDKLPHIFGFTLLGDFIYWTWQRRSIRVHVKKASRDVIIDQLPDLMLKAVNVA 597

Search completed: February 17, 2005, 01:36:50

Db 537 KTLLEDKLPHIFGFTLLGDFIYWTWQRRSIRVHVKKASRDVIIDQLPDLMLKAVNVA 596
Qy 598 KVGCTNPCADNRGCGSHLCPTPHATRCGCGIIGLELLSDMKTCIYPEAFIPTSAAATHR 657
Db 597 KVGCTNPCADNRGCGSHLCPTPHATRCGCGIIGLELLSDMKTCIYPEAFIPTSAAATHR 656
Qy 658 ISLETNNNDVAIPLTGVKEASALDPDVSNHHIYWTDSLKTISRAPFMGSSVEHVVEFGL 717
Db 657 ISLETNNNDVAIPLTGVKEASALDPDVSNHHIYWTDSLKTISRAPFMGSSVEHVVEFGL 716
Qy 718 DYPEGMAVDMMGKNLYNADTGTNRIEVARLDQGFQVLRDLDNPRSLALDPTKGYIYW 777
Db 717 DYPEGMAVDMMGKNLYNADTGTNRIEVARLDQGFQVLRDLDNPRSLALDPTKGYIYW 776
Qy 778 TEWGGKPRIVRAFMGDNCTMLVDKVGGRANDLTIDYADORLYWTDLTNMTSSNMLQGE 837
Db 777 TEWGGKPRIVRAFMGDNCTMLVDKVGGRANDLTIDYADORLYWTDLTNMTSSNMLQGE 836
Qy 838 RWIADDLPHFPGLTOYSDIYWTWDLNLSHISERADKTSGRNRTLIOGHLDVMDILVPHS 897
Db 837 RWIADDLPHFPGLTOYSDIYWTWDLNLSHISERADKTSGRNRTLIOGHLDVMDILVPHS 896
Qy 898 SRQDGLNDCHMNGCCGQCLAIPEGHRCGCASHYTLDPSNRNCSPTTFLFSSQKSAIS 957
Db 897 SRQDGLNDCHMNGCCGQCLAIPEGHRCGCASHYTLDPSNRNCSPTTFLFSSQKSAIS 956
Qy 958 RMIPODQHSPLIILPHGLRNKAIYDPLDKFIYWDGRQNIKRAKDDGTQPFVLTSL 1017
Db 957 RMIPODQHSPLIILPHGLRNKAIYDPLDKFIYWDGRQNIKRAKDDGTQPFVLTSL 1016
Qy 1018 QGONPDQPHDLSIDIYSRTLFTCEATNTINVRHLSGEAMGVLRGDRDKPRATVNAE 1077
Db 1017 QLSLSPDQPHDLSIDIYSRTLFTCEATNTINVRHLSGEAMGVLRGDRDKPRATVNAE 1076
Qy 1078 RGYLYFTNMQDRAAKIERASLDGTEREVLFTGLIRPVALVVDNLTGLKFWVDADLKRIE 1137
Db 1077 RGYLYFTNMQDRAAKIERASLDGTEREVLFTGLIRPVALVVDNLTGLKFWVDADLKRIE 1136
Qy 1138 SCDLSGANRLTLEDAIVQPLGLIILGKHLIWDROQMIERVEKTTGDKRTRIOGRAH 1197
Db 1137 SCDLSGANRLTLEDAIVQPLGLIILGKHLIWDROQMIERVEKTTGDKRTRIOGRAH 1196
Qy 1198 LTGTHAVEEVSLEBSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVQLNLLTCGSRPP 1257
Db 1197 LTGTHAVEEVSLEBSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVQLNLLTCGSRPP 1256
Qy 1258 TCSPDQFACATGEIDCIPGAWRCDFPECDQSDDEBGCPCVCSAAQFPCARGQCVDLRLC 1317
Db 1257 TCSPDQFACATGEIDCIPGAWRCDFPECDQSDDEBGCPCVCSAAQFPCARGQCVDLRLC 1316
Qy 1318 DGEADCDRSDRDAICLBNQRCASGQCVLIKQCCDSFDCIDGSDLMCEITKPPS 1377
Db 1317 DGEADCDRSDRDAICLBNQRCASGQCVLIKQCCDSFDCIDGSDLMCEITKPPS 1376
Qy 1378 DDPHSAISAIGPVIGIILSLFVMGVYVFCORVVCORVYAGANGPPPHYVSGTPHVLNLF 1437
Db 1377 DDPHSAISAIGPVIGIILSLFVMGVYVFCORVVCORVYAGANGPPPHYVSGTPHVLNLF 1436
Qy 1438 IAPGSGHQGPPTGIACGKSMMSVSLMGGRGVPLYDRNHVYTGASSSSSSSTKATLYPPI 1497
Db 1437 IAPGSGHQGPPTGIACGKSMMSVSLMGGRGVPLYDRNHVYTGASSSSSSSTKATLYPPI 1496
Qy 1498 LNPPSPATDPSLXNMDMFYSSNIPATAPRPRPIIRGMAPPTPCSDVCDSDYSASRW 1557
Db 1497 LNPPSPATDPSLXNMDMFYSSNIPATAPRPRPIIRGMAPPTPCSDVCDSDYSASRW 1556
Qy 1558 KASKYLDLNSDSDPYPPPTPHSHOYLSAEDSCPPSPATERSYPHLFPFPPSPCTDSS 1615
Db 1557 KASKYLDLNSDSDPYPPPTPHSHOYLSAEDSCPPSPATERSYPHLFPFPPSPCTDSS 1614

Job time : 58 secs

2221 GCCGTTGACTGGATGGGCAAGAACTCTACTGGGCCGACACTGGGACCAACAGAAATCGAA 2280
2221 GCCGTTGACTGGATGGGCAAGAACTCTACTGGGCCGACACTGGGACCAACAGAAATCGAA 2280
2281 GTGGCGCGGCTGGACGGGCAAGTTCCGGCAAGTCTCGTGTGGAGGACTTGGACAAACCG 2340
2281 GTGGCGCGGCTGGACGGGCAAGTTCCGGCAAGTCTCGTGTGGAGGACTTGGACAAACCG 2340
2341 AGGTGCGTGGCCCTGGATCCCAACAGGGCTACATCTACTGGACCGAGTGGGGGGCAAG 2400
2341 AGGTGCGTGGCCCTGGATCCCAACAGGGCTACATCTACTGGACCGAGTGGGGGGCAAG 2400
2401 CCGAGGATCGTGGGGCTTCAATGACCGGGACCAACTGATGACGCTGGTGGACAAAGGTG 2460
2401 CCGAGGATCGTGGGGCTTCAATGACCGGGACCAACTGATGACGCTGGTGGACAAAGGTG 2460
2461 GGCAGGATCGTGGGGCTTCAATGACCGGGACCAACTGATGACGCTGGTGGACAAAGGTG 2520
2461 GGCAGGATCGTGGGGCTTCAATGACCGGGACCAACTGATGACGCTGGTGGACAAAGGTG 2520
2521 GACACCAACATGATCGAGTCTCGGCTTCAACATGCTGGGTGAGGACGGGTCTGATTCGGAC 2580
2521 GACACCAACATGATCGAGTCTCGGCTTCAACATGCTGGGTGAGGACGGGTCTGATTCGGAC 2580
2581 GATCTCCCGCACCCGTTCCGCTTCAACATGCTGGGTGAGGACGGGTCTGATTCGGAC 2640
2581 GATCTCCCGCACCCGTTCCGCTTCAACATGCTGGGTGAGGACGGGTCTGATTCGGAC 2640
2641 AATCTGCAACAGATTTGAGCGGCGGCGACAAAGACTAGCGCGCGGAAACCGCAACCTCATCCAG 2700
2641 AATCTGCAACAGATTTGAGCGGCGGCGACAAAGACTAGCGCGCGGAAACCGCAACCTCATCCAG 2700
2701 GGCACCTGGACTTGGATGAGACATCTGGTGTTCACATCTCTCCGCGACGATGGCCTC 2760
2701 GGCACCTGGACTTGGATGAGACATCTGGTGTTCACATCTCTCCGCGACGATGGCCTC 2760
2761 AATGACTGTATGCAACAAAGGGCAGTGTGGGACGTGTGCTTGGCCATCCCGGGCGGC 2820
2761 AATGACTGTATGCAACAAAGGGCAGTGTGGGACGTGTGCTTGGCCATCCCGGGCGGC 2820
2821 CACCGCTGGGCTGGCTTCACTACATACACCTTGGACCCGACGAGCGGCAATGACGACCGC 2880
2821 CACCGCTGGGCTGGCTTCACTACATACACCTTGGACCCGACGAGCGGCAATGACGACCGC 2880
2881 CCCACCACTTCTGCTTGTGAGCCAGAAATCTGCCATCAGTCCGATGATCCCGGACGAC 2940
2881 CCCACCACTTCTGCTTGTGAGCCAGAAATCTGCCATCAGTCCGATGATCCCGGACGAC 2940
2941 CAGCACAGCCCGGATCTCATCTGCGCTGTGATGAGTGTGAGGAACTGAGGAACTGAGGAACTGAG 3000
2941 CAGCACAGCCCGGATCTCATCTGCGCTGTGATGAGTGTGAGGAACTGAGGAACTGAGGAACTGAG 3000
3001 TATGACCACTTGGACAAAGTTCACTACTGAGTGGTGGATGGGCGCCAGAAATCAAGCGAGCC 3060
3001 TATGACCACTTGGACAAAGTTCACTACTGAGTGGTGGATGGGCGCCAGAAATCAAGCGAGCC 3060
3061 AAGGACGAGGGGACCCAGCCCTTGTGTTGACCTCTCTGAGCCAAAGGCGGCAAAACCCAGAC 3120
3061 AAGGACGAGGGGACCCAGCCCTTGTGTTGACCTCTCTGAGCCAAAGGCGGCAAAACCCAGAC 3120
3121 AAGGACGAGGGGACCCAGCTCAGCATCGACATCTACAGCCGGAACACTGTTCTGACGTCGAG 3180
3121 AAGGACGAGGGGACCCAGCTCAGCATCGACATCTACAGCCGGAACACTGTTCTGACGTCGAG 3180
3181 GCCACCAATACATCAACGTTCCAGAGTGGACGCGGGGAAAGCAATGGGGGTGGTCTGCGT 3240
3181 GCCACCAATACATCAACGTTCCAGAGTGGACGCGGGGAAAGCAATGGGGGTGGTCTGCGT 3240
3241 GGGGACCGGACNAGCCAGGGCCATCTGTCAGCGGAGCGAGGGGTACTGTTACTTC 3300
3241 GGGGACCGGACNAGCCAGGGCCATCTGTCAGCGGAGCGAGGGGTACTGTTACTTC 3300

3301 ACCAACATGACGAGCAGCGGCGAGCCAAAGATCGAAACGCGAGCCCTGGACGGCAACGAGCGC 3360
3301 ACCAACATGACGAGCAGCGGCGAGCCAAAGATCGAAACGCGAGCCCTGGACGGCAACGAGCGC 3360
3361 GAGGTCTCTTTCAACAACGCGGCTCATCCGCCCTGTGGCCCTGTGGTGGAGCAACACACTG 3420
3361 GAGGTCTCTTTCAACAACGCGGCTCATCCGCCCTGTGGCCCTGTGGTGGAGCAACACACTG 3420
3421 GGCAGGCTGTTCTGGGTGGACGCGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTGACGG 3480
3421 GGCAGGCTGTTCTGGGTGGACGCGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTGACGG 3480
3481 GCCAACCGCTGACCCCTGGAGAGCGCCAAACATCTGTCAGCCTCTGGGGCTGACCATCTTT 3540
3481 GCCAACCGCTGACCCCTGGAGAGCGCCAAACATCTGTCAGCCTCTGGGGCTGACCATCTTT 3540
3541 GGCAGGCTGTTCTGGGTGGACGCGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTGACGG 3600
3541 GGCAGGCTGTTCTGGGTGGACGCGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTGACGG 3600
3601 ACCGGGGAACGAGCTGCGATCCAGGGCGTGTGGCCCACTTCACCTGAGCTGATCCATGCA 3660
3601 ACCGGGGAACGAGCTGCGATCCAGGGCGTGTGGCCCACTTCACCTGAGCTGATCCATGCA 3660
3661 GTGAGGAAAGTCAAGCTGAGGAGTCTCAGCCCAACCATGTGCCCTGACAAATGGTGGC 3720
3661 GTGAGGAAAGTCAAGCTGAGGAGTCTCAGCCCAACCATGTGCCCTGACAAATGGTGGC 3720
3721 TGCTCCCAACATCTGTATTTGCCAAGGTGATGGGACACAACCGGTGCTCATGCCCAGTCCAC 3780
3721 TGCTCCCAACATCTGTATTTGCCAAGGTGATGGGACACAACCGGTGCTCATGCCCAGTCCAC 3780
3781 CTCGTGCTCTGCAAGAACCTGCTGACCTGTGGAGAGCGCCCACTGCTCCCGGACGAG 3840
3781 CTCGTGCTCTGCAAGAACCTGCTGACCTGTGGAGAGCGCCCACTGCTCCCGGACGAG 3840
3841 TTTGCATGTGCAACAGGGGAGATCGACTGTATCCCGGGGCTGGCGCTGTGACCGGCTTT 3900
3841 TTTGCATGTGCAACAGGGGAGATCGACTGTATCCCGGGGCTGGCGCTGTGACCGGCTTT 3900
3901 CCCGAGTGCATGATGACAGAGCGAGAGGAGTGCCTCCGTTGCTCCGCGCCCGAGTTC 3960
3901 CCCGAGTGCATGATGACAGAGCGAGAGGAGTGCCTCCGTTGCTCCGCGCCCGAGTTC 3960
3961 CCTGCGCGGGGCTCAGTGTGGACCTGCGCTGCGCTGCGAGCGGCGAGGACGACTGT 4020
3961 CCTGCGCGGGGCTCAGTGTGGACCTGCGCTGCGCTGCGAGCGGCGAGGACGACTGT 4020
4021 CAGGACCGCTCAGACGAGGCGGACTGTGACGCCATCTGCTGCGCCCAACAGTTCGGGTGT 4080
4021 CAGGACCGCTCAGACGAGGCGGACTGTGACGCCATCTGCTGCGCCCAACAGTTCGGGTGT 4080
4081 GCGAGCGGCGAGTGTGCTCATCAACAGAGTGGGACTCCTTCCCGGACTGTATCGAC 4140
4081 GCGAGCGGCGAGTGTGCTCATCAACAGAGTGGGACTCCTTCCCGGACTGTATCGAC 4140
4141 GGCCTCCGACGAGCTCATGTGTGAATCAACAGCGGCGCTCAGAGCAGCGCGGCCAC 4200
4141 GGCCTCCGACGAGCTCATGTGTGAATCAACAGCGGCGCTCAGAGCAGCGCGGCCAC 4200
4201 AGCAGTGCATCGGCGGCTCATTTGGCATCATCTCTCTCTTCGTTCATGGGTGGTTC 4260
4201 AGCAGTGCATCGGCGGCTCATTTGGCATCATCTCTCTCTTCGTTCATGGGTGGTTC 4260
4261 TATTTTGTGTGCCAGCGGTGTGTCAGCGCTATGCGGGGGCCAAAGCGGCGCTTCGG 4320
4261 TATTTTGTGTGCCAGCGGTGTGTCAGCGCTATGCGGGGGCCAAAGCGGCGCTTCGG 4320
4321 CACGAGTATGTACGCGGAGCCCGCAGTGCCTCTCAATTTATAGCCCGCGGGGTTC 4380
4321 CACGAGTATGTACGCGGAGCCCGCAGTGCCTCTCAATTTATAGCCCGCGGGGTTC 4380
4381 CAGCATGGCCCTTTACAGGCGATCGCATGCGGAAAGTCCATGATGAGCTCGGTGAGCGCTG 4440

Db	4381		CAGCATGGCCCTTTCACAGGCATCGCATGCGGAAGTTCATGATGACTCCGTGAGCGTG	4440
Qy	4441		ATGGGGGCGGGCGGGGTGCCCTTGTCAGACCGGAACCAACGCTCACAGGGGGCTTCGTCC	4500
Db	4441		ATGGGGGCGGGCGGGGTGCCCTTGTCAGACCGGAACCAACGCTCACAGGGGGCTTCGTCC	4500
Qy	4501		AGCAGCTCGTCCACGACGAGAGGCCACACCTGTACCGCGGATCTCTGAACCCGCGCCCTCC	4560
Db	4501		AGCAGCTCGTCCACGACGAGAGGCCACACCTGTACCGCGGATCTCTGAACCCGCGCCCTCC	4560
Qy	4561		CCGGCCACGGACCCCTCCCTGTGTAAACATGACATGTTCTTACTCTTCAAAACATTCGGGCC	4620
Db	4561		CCGGCCACGGACCCCTCCCTGTGTAAACATGACATGTTCTTACTCTTCAAAACATTCGGGCC	4620
Qy	4621		ACTCGGAGCCGTACAGGCCCTACATCATTCGAGGAATGGGCCCCCGACGAGCCCTGTC	4680
Db	4621		ACTCGGAGCCGTACAGGCCCTACATCATTCGAGGAATGGGCCCCCGACGAGCCCTGTC	4680
Qy	4681		AGCACCGACGTGTGTGACAGCACTACAGGGCCAGCCGCTGTGAAGGCCACCAAGTACTAC	4740
Db	4681		AGCACCGACGTGTGTGACAGCACTACAGGGCCAGCCGCTGTGAAGGCCACCAAGTACTAC	4740
Qy	4741		CTGGATTGTGAACCTGGACTCAGACCCCTATCCACCCCGACCCAGCCGCCACGACGATAC	4800
Db	4741		CTGGATTGTGAACCTGGACTCAGACCCCTATCCACCCCGACCCAGCCGCCACGACGATAC	4800
Qy	4801		CTGTCGGCGGAGACAGCTGCCGCCCTCGCCGCCACCGAGAGGAGCTACTTCCATCTC	4860
Db	4801		CTGTCGGCGGAGACAGCTGCCGCCCTCGCCGCCACCGAGAGGAGCTACTTCCATCTC	4860
Qy	4861		TTCCGGCCCTCGTCCCTCGACGGACTCATCTGACCTCGCGGGGCCACTCTGGC	4920
Db	4861		TTCCGGCCCTCGTCCCTCGACGGACTCATCTGACCTCGCGGGGCCACTCTGGC	4920
Qy	4921		TTCTCTGTGCCCTGTAAATAGTTTTAAATATGAACAAAGAAAAAATATATTTTATGAT	4980
Db	4921		TTCTCTGTGCCCTGTAAATAGTTTTAAATATGAACAAAGAAAAAATATATTTTATGAT	4980
Qy	4981		TTAAAAATAAATAATATCGGATTTTAAAAACATGAGAAATCTGAACTGTGATGGGTG	5040
Db	4981		TTAAAAATAAATAATATCGGATTTTAAAAACATGAGAAATCTGAACTGTGATGGGTG	5040
Qy	5041		GGCAGGCTCGGAGAACTTTGTA	5063
Db	5041		GGCAGGCTCGGAGAACTTTGTA	5063

RESULT 2

```

US-10-374-979-1
; Sequence 1, Application No.10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 1
; LENGTH: 5120
; TYPE: DNA
; ORGANISM: Homo sapiens

```

QY	1021	CCCACGGGTGTGACGTGACGACCAACGCGAGACGCTGTAAAGCAGGAGCCGAGGAGGTG	1080
Db	1040	CCCACGGGTGTGACGTGACGACCAACGCGAGACGCTGTAAAGCAGGAGCCGAGGAGGTG	1099
QY	1081	CTGCTGCTGGCCGCGCGAGACGGAACCTAAGAGAGTCTCGCTGGACACGCCCGGACTTCACC	1140
Db	1100	CTGCTGCTGGCCGCGCGAGACGGAACCTAAGAGAGTCTCGCTGGACACGCCCGGACTTCACC	1159
QY	1141	GACATCGTGTGCTGAGTGGAGACATCCGGACCGCATTTGCCATCGACTACGACCCGCTA	1200
Db	1160	GACATCGTGTGCTGAGTGGAGACATCCGGACCGCATTTGCCATCGACTACGACCCGCTA	1219
QY	1201	GAGGCTATGTCTACTTGCACAGATGACGAGGTGCGGGCCATCCGAGGCGGTACCTGGAC	1260
Db	1220	GAGGCTATGTCTACTTGCACAGATGACGAGGTGCGGGCCATCCGAGGCGGTACCTGGAC	1279
QY	1261	GGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCCGCGTC	1320
Db	1280	GGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCCGCGTC	1339
QY	1321	GACTGGGTGGCCGGAACCTCTACTGGACCGACAGCGGACCGGATCGAGGTGACG	1380
Db	1340	GACTGGGTGGCCGGAACCTCTACTGGACCGGACAGCGGATCGAGGTGACG	1399
QY	1381	CGCTCTAACCGGACCTCCCGCAAGATCTGTGTGCGAGGACCTGAGACGAGCCCGAGGCC	1440
Db	1400	CGCTCTAACCGGACCTCCCGCAAGATCTGTGTGCGAGGACCTGAGACGAGCCCGAGGCC	1459
QY	1441	ATCGCACTGCAACCCCGTGTATGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAAA	1500
Db	1460	ATCGCACTGCAACCCCGTGTATGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAAA	1519
QY	1501	ATCGAGTGTGCAACTTGGATGGGACGAGCGGCGTGTGCTGATCAATGCCTCCCTCGGG	1560
Db	1520	ATCGAGTGTGCAACTTGGATGGGACGAGCGGCGTGTGCTGATCAATGCCTCCCTCGGG	1579
QY	1561	TGGCCCAACCGCCTGGACCTGCAAGAGGGAAGCTCTACTGGGGAGAGCCCAAG	1620
Db	1580	TGGCCCAACCGCCTGGACCTGCAAGAGGGAAGCTCTACTGGGGAGAGCCCAAG	1639
QY	1621	ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCTCTCGAGGAC	1680
Db	1640	ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCTCTCGAGGAC	1699
QY	1681	AAGCTCCCGCAATTTTCGGGTTCACGCTGTGCGGGACCTTCATCTACTGACTGACTGG	1740
Db	1700	AAGCTCCCGCAATTTTCGGGTTCACGCTGTGCGGGACCTTCATCTACTGACTGACTGG	1759
QY	1741	CAGCGCCGACATCGAGCGGTGCAAGAGTCAAGCCAGCGGAGCTCATCATTTGAC	1800
Db	1760	CAGCGCCGACATCGAGCGGTGCAAGAGTCAAGCCAGCGGAGCTCATCATTTGAC	1819
QY	1801	CAGCTGCCCGACCTGATGGGCTCAAAGCTGTGAATGTGGCCAAAGTCTGCGAACCAC	1860
Db	1820	CAGCTGCCCGACCTGATGGGCTCAAAGCTGTGAATGTGGCCAAAGTCTGCGAACCAC	1879
QY	1861	CCGTGTGCGGACAGGAACGGGGGTGACGACCTGTGCTTTCACACCCCGACCAACC	1920
Db	1880	CCGTGTGCGGACAGGAACGGGGGTGACGACCTGTGCTTTCACACCCCGACCAACC	1939
QY	1921	CGGTGTGCTGCCCATCGGCTGAGCTGTGATGACATGAGACCTGATCGTGCCT	1980
Db	1940	CGGTGTGCTGCCCATCGGCTGAGCTGTGATGACATGAGACCTGATCGTGCCT	1999
QY	1981	GAGGCTTTCTTTGGTCTTCAACAGCAGAGCGCGCATCCACAGGATCTCCCTCGAGACCAAT	2040
Db	2000	GAGGCTTTCTTTGGTCTTCAACAGCAGAGCGCGCATCCACAGGATCTCCCTCGAGACCAAT	2059
QY	2041	AACAAAGAGTGGCCATCCCGTCAAGGGCGTCAAGGAGGCGCTCAGCCCTGGACTTTGAT	2100
Db	2060	AACAAAGAGTGGCCATCCCGTCAAGGGCGTCAAGGAGGCGCTCAGCCCTGGACTTTGAT	2119
QY	2101	GTGTCCCAACCAACCATCTACTGGACAGACGCTGACCGCTGAAGACCATCAGCGCGCCTTC	2160
Db	2120	GTGTCCCAACCAACCATCTACTGGACAGACGCTGACCGCTGAAGACCATCAGCGCGCCTTC	2179
QY	2161	ATGAAACGGGAGCTCGGTGGAGCAAGTGTGTGAGTTTGGCTTGAATAACCGAGGGCATG	2220
Db	2180	ATGAAACGGGAGCTCGGTGGAGCAAGTGTGTGAGTTTGGCTTGAATAACCGAGGGCATG	2239
QY	2221	GCCGTGTGATCTGATGGGCAAGAACTCTACTTGTGGCCGACACTTGGGACCAACAGATTCGA	2280
Db	2240	GCCGTGTGATCTGATGGGCAAGAACTCTACTTGTGGCCGACACTTGGGACCAACAGATTCGA	2299
QY	2281	GTGGCGGCTTGGACGGGAGTTCCGGCAAGTCTCTGTGTGGAGGACTTGGACAAACCG	2340
Db	2300	GTGGCGGCTTGGACGGGAGTTCCGGCAAGTCTCTGTGTGGAGGACTTGGACAAACCG	2359
QY	2341	AGTCTGTGGCCCTGGATCCCAAGGGCTTACTACTTGAACCGAGTGGGGCGGCAAG	2400
Db	2360	AGTCTGTGGCCCTGGATCCCAAGGGCTTACTACTTGAACCGAGTGGGGCGGCAAG	2419
QY	2401	CCGAGGATCGTGGCGGCTTCATGGACGGGACCAACTGCATGACGCTGGTGGACAGGTG	2460
Db	2420	CCGAGGATCGTGGCGGCTTCATGGACGGGACCAACTGCATGACGCTGGTGGACAGGTG	2479
QY	2461	GGCGGGCAACGACCTCACCAATTGACTACGCTGACACGCGCTTACTTGGACCGACCTG	2520
Db	2480	GGCGGGCAACGACCTCACCAATTGACTACGCTGACACGCGCTTACTTGGACCGACCTG	2539
QY	2521	GACACCAACATGATCGAGTCTGCAACATGCTGGGTGAGGACGGGTCTGATTCGCGAC	2580
Db	2540	GACACCAACATGATCGAGTCTGCAACATGCTGGGTGAGGACGGGTCTGATTCGCGAC	2599
QY	2581	GATCTCCGCGACCCGTTTGGTCTGACCGGAGTACAGCGATATATCTACTTGGACAGCTGG	2640
Db	2600	GATCTCCGCGACCCGTTTGGTCTGACCGGAGTACAGCGATATATCTACTTGGACAGCTGG	2659
QY	2641	AATCTGCACAGCATTTGACGGCGCGCAAGAACTAGCGCCGGAACCGCACTCATCCAG	2700
Db	2660	AATCTGCACAGCATTTGACGGCGCGCAAGAACTAGCGCCGGAACCGCACTCATCCAG	2719
QY	2701	GGCAACCTGTGATGAGTGAACATCTCTGTGTTCATCTCTCCCGCAGGATGGCTTC	2760
Db	2720	GGCAACCTGTGATGAGTGAACATCTCTGTGTTCATCTCTCCCGCAGGATGGCTTC	2779
QY	2761	AATGACTGTATGCAACAAACGGGCGAGTGTGGGAGCTGTGCTTCCATCCCGCGCGC	2820
Db	2780	AATGACTGTATGCAACAAACGGGCGAGTGTGGGAGCTGTGCTTCCATCCCGCGCGC	2839
QY	2821	CACCGCTGCGGCTCGGCTCACACTACACCTTGACCCCGCAGCGCCAACTGCAAGCCCG	2880
Db	2840	CACCGCTGCGGCTCGGCTCACACTACACCTTGACCCCGCAGCGCCAACTGCAAGCCCG	2899
QY	2881	CCCAACCACTTTCTGTGCTTTCAGCCAGAAATCTGCCATCAGTCGGATGATCCCGGACGAC	2940
Db	2900	CCCAACCACTTTCTGTGCTTTCAGCCAGAAATCTGCCATCAGTCGGATGATCCCGGACGAC	2959
QY	2941	CAGCAGCCCGGATCTCATCTGCTCCCTGATGGAAGTCTGAGGAGCTCAAGACCAATCGAC	3000
Db	2960	CAGCAGCCCGGATCTCATCTGCTCCCTGATGGAAGTCTGAGGAGCTCAAGACCAATCGAC	3019
QY	3001	TATGACCCACCTGGACAAAGTTTCATCTACTGGGTGATGGGCGCCAGAAATCAAGCGAGCC	3060
Db	3020	TATGACCCACCTGGACAAAGTTTCATCTACTGGGTGATGGGCGCCAGAAATCAAGCGAGCC	3079
QY	3061	AAGAACGAGGAGCCGAGCCCTTTTGTGACTCTCTGAGCCCAAGGCAAAACCCAGAC	3120
Db	3080	AAGAACGAGGAGCCGAGCCCTTTTGTGACTCTCTGAGCCCAAGGCAAAACCCAGAC	3139
QY	3121	AGGACGAGGAGCCGAGCCCTGAGCATCGACATCTTACAGCGGACACTGTTCTGACGTCGAG	3180
Db	3140	AGGACGAGGAGCCGAGCCCTGAGCATCGACATCTTACAGCGGACACTGTTCTGACGTCGAG	3199
QY	3181	GCCACCAATACCATCAACGCTCCACAGGCTGAGCGGGGAGGCAATGGGGGTGTGTGCTGCT	3240

|||||
1940 CGGTGTGGCTGCCCATCGGCTGGAGCTGCTGAGTGACATGAAGACCTGATCGTGCCT 1999
QY
1981 GAGGCTTTCTTGGTCTTTACACAGCAGAGCGGCCATCCACAGGATCTCCCTCGAGACCAAT 2040
Db
2000 GAGGCTTTCTTGGTCTTTACACAGCAGAGCGGCCATCCACAGGATCTCCCTCGAGACCAAT 2059
QY
2041 AACAAAGAGCTGGGCTATCCGCTCAACGGGCTCAAGGAGGCTCAGCCCTGGACTTTGAT 2100
Db
2060 AACAAAGAGCTGGGCTATCCGCTCAACGGGCTCAAGGAGGCTCAGCCCTGGACTTTGAT 2119
QY
2101 GTGTCCAAACCAACATCTACTGACACAGCTCAGCTGGAAGACCATCAGCCGCGCTTC 2160
Db
2120 GTGTCCAAACCAACATCTACTGACACAGCTCAGCTGGAAGACCATCAGCCGCGCTTC 2179
QY
2161 ATGAACGGGAGCTCGGTGGAGCAAGTGTGTGAGTTTGGCTTTGACTACCCGAGGGGATG 2220
Db
2180 ATGAACGGGAGCTCGGTGGAGCAAGTGTGTGAGTTTGGCTTTGACTACCCGAGGGGATG 2239
QY
2221 GCCGTTGACTGGATGGGCAAGAACTTACTTGGGCGGACACTGGAGCCCAACAGAAATCGAA 2280
Db
2240 GCCGTTGACTGGATGGGCAAGAACTTACTTGGGCGGACACTGGAGCCCAACAGAAATCGAA 2299
QY
2281 GTGGCGGCTGGAGCGGAGCTTCGGCAAGTCTCTGTTGGAGGAGCTTGGACAAACCCG 2340
Db
2300 GTGGCGGCTGGAGCGGAGCTTCGGCAAGTCTCTGTTGGAGGAGCTTGGACAAACCCG 2359
QY
2341 AGGTGCTGGGCTGGATGCCCAACCAAGGGCTTACTTACTGGAGCGAGTGGGGGGCGGCAAG 2400
Db
2360 AGGTGCTGGGCTGGATGCCCAACCAAGGGCTTACTTACTGGAGCGAGTGGGGGGCGGCAAG 2419
QY
2401 CCGAGGATCGTGGGCTTCTATGGAAGGAGCACTGATGACGCTGCTGAGCAAGGTG 2460
Db
2420 CCGAGGATCGTGGGCTTCTATGGAAGGAGCACTGATGACGCTGCTGAGCAAGGTG 2479
QY
2461 GGCGGGGCAACGAGCTCACATTTGACTTACGCTGACAGCGCTTCTACTGGACCGACCTG 2520
Db
2480 GGCGGGGCAACGAGCTCACATTTGACTTACGCTGACAGCGCTTCTACTGGACCGACCTG 2539
QY
2521 GACACCAACATGATCGAGTGTGTCACCAATGTGGGTGAGGAGCGGCTGATGATGCCGAC 2580
Db
2540 GACACCAACATGATCGAGTGTGTCACCAATGTGGGTGAGGAGCGGCTGATGATGCCGAC 2599
QY
2581 GATCTCCGCAACCGCTTCTGCTGACGAGTACAGGATTTATCTACTTGGACAGACTGG 2640
Db
2600 GATCTCCGCAACCGCTTCTGCTGACGAGTACAGGATTTATCTACTTGGACAGACTGG 2659
QY
2641 AATCTGCACAGCATTTGAGCGGCGGACCAAGACTAGCGGCGGAAACCGCAACCTCATCCAG 2700
Db
2660 AATCTGCACAGCATTTGAGCGGCGGACCAAGACTAGCGGCGGAAACCGCAACCTCATCCAG 2719
QY
2701 GGCGACCTGGATCTGATGAGACATCTGTTGTTTCACTCTCTCCGCGAGGATGGCTTC 2760
Db
2720 GGCGACCTGGATCTGATGAGACATCTGTTGTTTCACTCTCTCCGCGAGGATGGCTTC 2779
QY
2761 AATGACTGATGACAAACAAAGGAGTGTGGGAGCTGTGCTTGGCAATCCCGCGGCGC 2820
Db
2780 AATGACTGATGACAAACAAAGGAGTGTGGGAGCTGTGCTTGGCAATCCCGCGGCGC 2839
QY
2821 CACCGCTGCGGCTGGCTTCACTACACCTTGGACCCCGAGCAGCGCAACTGACAGCCG 2880
Db
2840 CACCGCTGCGGCTGGCTTCACTACACCTTGGACCCCGAGCAGCGCAACTGACAGCCG 2899
QY
2881 CCCACGACTTCTGTTGTTGAGCAGAAATCTGCGCATCTGATGATGATCCGAGCGAC 2940
Db
2900 CCCACGACTTCTGTTGTTGAGCAGAAATCTGCGCATCTGATGATGATCCGAGCGAC 2959
QY
2941 CAGCAGCGCGGATCTCATCTCTGCTGCTGACTGAGGAAAGTCAAAAGCCATCGAC 3000
Db
2960 CAGCAGCGCGGATCTCATCTCTGCTGCTGCTGACTGAGGAAAGTCAAAAGCCATCGAC 3019
QY
3001 TATGACCCACTGGACAAAGTTTCACTGTTGGTGGATGGGCGCCAGAACTTACGAGCGGC 3060
|||||

Db
3020 TATGACCCACTGGACAAAGTTTCACTGTTGGTGGATGGGCGCCAGAACATCAAGCGAGCC 3079
QY
3061 AAGGACGAGGGGACCCAGCCCTTGTGTTGACCTCTCTGAGCCAAAGCCAAAACCCAGAC 3120
Db
3080 AAGGACGAGGGGACCCAGCCCTTGTGTTGACCTCTCTGAGCCAAAGCCAAAACCCAGAC 3139
QY
3121 AGGACAGCCCAACCACTCAGCATCGACATCTACAGCCGAGACACTGTTCTGGAGGTGCGAG 3180
Db
3140 AGGACAGCCCAACCACTCAGCATCGACATCTACAGCCGAGACACTGTTCTGGAGGTGCGAG 3199
QY
3181 GCCAACAAATACCATCAAGCTCCAGGCTGAGCGGGGAAAGCCATGGGGGTGTTGCTCGT 3240
Db
3200 GCCAACAAATACCATCAAGCTCCAGGCTGAGCGGGGAAAGCCATGGGGGTGTTGCTCGT 3259
QY
3241 GGGGACCGGACAAAGCCAGGGGCTCATCGTCTCAAGCGGAGGAGGTACCTGTACTTTC 3300
Db
3260 GGGGACCGGACAAAGCCAGGGGCTCATCGTCTCAAGCGGAGGAGGTACCTGTACTTTC 3319
QY
3301 ACCAAACATGAGGACCGGGGAGCCAAAGATCGAAGCGGACGCTTGGACGGACCGAGCGC 3360
Db
3320 ACCAAACATGAGGACCGGGGAGCCAAAGATCGAAGCGGACGCTTGGACGGACCGAGCGC 3379
QY
3361 GAGTCTCTTTCACCAACCGGCTCATCCGCCCTGTGCGCTTGTGGGTGGACCAACACTG 3420
Db
3380 GAGTCTCTTTCACCAACCGGCTCATCCGCCCTGTGCGCTTGTGGGTGGACCAACACTG 3439
QY
3421 GGCAGGCTGTTTGGGTGGAGCGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTCAAGG 3480
Db
3440 GGCAGGCTGTTTGGGTGGAGCGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTCAAGG 3499
QY
3481 GCCAACCGCTGACCTTGGAGGAGCGCAACATCGTGCAGCTCTTGGGCTTGACCATCTT 3540
Db
3500 GCCAACCGCTGACCTTGGAGGAGCGCAACATCGTGCAGCTCTTGGGCTTGACCATCTT 3559
QY
3541 GGCAGGATCTTACTGATGAGCCGCGGACGAGAGATGATCGAGCTGTGGAGAAAGAC 3600
Db
3560 GGCAGGATCTTACTGATGAGCCGCGGACGAGAGATGATCGAGCTGTGGAGAAAGAC 3619
QY
3601 ACCGGGAGCAAGCGGACTCGCATCCAGGGCGCTGCGCCACCTCACTGGCATCATGCA 3660
Db
3620 ACCGGGAGCAAGCGGACTCGCATCCAGGGCGCTGCGCCACCTCACTGGCATCATGCA 3679
QY
3661 GTGAGGAAAGTCAAGCTTGGAGGAGTTCTAGCCCAACCATGTGCGGCTGAGCAATGTGGC 3720
Db
3680 GTGAGGAAAGTCAAGCTTGGAGGAGTTCTAGCCCAACCATGTGCGGCTGAGCAATGTGGC 3739
QY
3721 TGCTCCCAACATCTGATTTGCCAGGAGTATGAGGACACACGCTGCTCATGCCAGTCCAC 3780
Db
3740 TGCTCCCAACATCTGATTTGCCAGGAGTATGAGGACACACGCTGCTCATGCCAGTCCAC 3799
QY
3781 CTCGTGCTCTTCGAGAAACCTGCTGACCTGTGGAGAGCGGCCACCTGCTCCCCGAGACAG 3840
Db
3800 CTCGTGCTCTTCGAGAAACCTGCTGACCTGTGGAGAGCGGCCACCTGCTCCCCGAGACAG 3859
QY
3841 TTTGATGTGCAACAGGGGAGATCGACTGTATATCCCGGGGCTGGCGCTGTGACGGCTTT 3900
Db
3860 TTTGATGTGCAACAGGGGAGATCGACTGTATATCCCGGGGCTGGCGCTGTGACGGCTTT 3919
QY
3901 CCCGAGTGCAGTACAGAGCGAGAGGAGGCTGCGCGCTGTCTCCCGCCCGAGTTTC 3960
Db
3920 CCCGAGTGCAGTACAGAGCGAGAGGAGGCTGCGCGCTGTCTCCCGCCCGAGTTTC 3979
QY
3961 CCTGCGCGCGGGGTCAAGTGTGGACCTGCGCTTGGCTGCGAGCGGAGGAGAGTGT 4020
Db
3980 CCTGCGCGCGGGGTCAAGTGTGGACCTGCGCTTGGCTGCGAGCGGAGGAGAGTGT 4039
QY
4021 CAGGACCGCTCAGAGCGGAGCTGTGACAGGCACTGCTGCTGCGCCCAACAGTTCGGGT 4080
Db
4040 CAGGACCGCTCAGAGCGGAGGAGTGTGACAGGCACTGCTGCTGCGCCCAACAGTTCGGGT 4099
QY
4081 CCGAGCGGCGAGTGTCTCATCAACAGAGTGGAGTCTCTTCCCCGAGTGTATCGAC 4140
Db
4100 CCGAGCGGCGAGTGTCTCATCAACAGAGTGGAGTCTCTTCCCCGAGTGTATCGAC 4159
pb

Qy	721	CTGGACGGCTCTTCCGSCAGAGTGTGGAGGSCAGCTGACGCAACCCCTTCGCCCTG	780	1801	CAGCTGCCCGACCTGATGGGGCTCAAAGCTGTGAATGTGGCCAAAGGTCTGTGGAAACCAAC	1860
Db	740	CTGGACGGCTCTTCCGSCAGAGTGTGGAGGSCAGCTGACGCAACCCCTTCGCCCTG	799	1820	CAGTGTCCCGACCTGATGGGGCTCAAAGCTGTGAATGTGGCCAAAGGTCTGTGGAAACCAAC	1879
Qy	781	ACGCTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCCGCTCCATCCATGCGCTGC	840	1861	CCGTGTCCGGAAGAGAAACGGGGGTGAGCACTGTGTCTTTTCAACCCCAACCAAC	1920
Db	800	ACGCTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCCGCTCCATCCATGCGCTGC	859	1880	CCGTGTCCGGAAGAGAAACGGGGGTGAGCACTGTGTCTTTTCAACCCCAACCAAC	1939
Qy	841	AACAAAGCGCACTGGGGGAGAGAGAGATCTCTGAGTGCCTCTACTCAACCAATGGAC	900	1921	CGTGTGGCTGCCCATCGGCTCGAGCTGTGATGACATGAAGACTGTGATCTGTGCT	1980
Db	860	AACAAAGCGCACTGGGGGAGAGAGAGATCTCTGAGTGCCTCTACTCAACCAATGGAC	919	1940	CGTGTGGCTGCCCATCGGCTCGAGCTGTGATGACATGAAGACTGTGATCTGTGCT	1999
Qy	901	ATCCAGGTGTGAGCCAGGAGCGGCTTTCTTCCACACTCGCTGTGAGGAGCAAT	960	1981	GAGGCTTCTTGTGCTTCCAGCAGAGCGGCCATCCACAGATCTCCCTCGAGACCAAT	2040
Db	920	ATCCAGGTGTGAGCCAGGAGCGGCTTTCTTCCACACTCGCTGTGAGGAGCAAT	979	2000	GAGGCTTCTTGTGCTTCCAGCAGAGCGGCCATCCACAGATCTCCCTCGAGACCAAT	2059
Qy	961	GGCGCTGCTCCCACTGTGCTGTCTCCCAAGCGAGCTTTCTACATGCGCTGC	1020	2041	AACAAAGCGCTGCGCTTCCCGCTCAAGGAGGCTCAGGCTCGGCTGTGAT	2100
Db	980	GGCGCTGCTCCCACTGTGCTGTCTCCCAAGCGAGCTTTCTACATGCGCTGC	1039	2060	AACAAAGCGCTGCGCTTCCCGCTCAAGGAGGCTCAGGCTCGGCTGTGAT	2119
Qy	1021	CCCAAGGTGTGAGCTGCAGCAAAAGGAGAGCTGTAAAGGAGAGCGGAGAGGTG	1080	2101	GTGTGCAACCAACATCTACTGGAAGAGCTGAGAGCAATCAGCGCGCTTC	2160
Db	1040	CCCAAGGTGTGAGCTGCAGCAAAAGGAGAGCTGTAAAGGAGAGCGGAGAGGTG	1099	2120	GTGTGCAACCAACATCTACTGGAAGAGCTGAGAGCAATCAGCGCGCTTC	2179
Qy	1081	CTGCTGTGCGCCGCGAGCGGACTACGAGAGATCTCGCTGGACAGCGCGGCTTCAAC	1140	2161	ATGAAACGGGAGCTCGGTGGAGCACTGTGAGTGTGGCTTGTACTACCCGAGGCGATG	2220
Db	1100	CTGCTGTGCGCCGCGAGCGGACTACGAGAGATCTCGCTGGACAGCGCGGCTTCAAC	1159	2180	ATGAAACGGGAGCTCGGTGGAGCACTGTGAGTGTGGCTTGTACTACCCGAGGCGATG	2239
Qy	1141	GACATCGTGTGCAAGTGGACGACATCCGCGACGCCATTTGCCATCGATACGACCCGCTA	1200	2221	GCCCTTGTGATGGAGGGAAGAACTCTACTGTGGGCGGACACTGGGACCAACAGATCGAA	2280
Db	1160	GACATCGTGTGCAAGTGGACGACATCCGCGACGCCATTTGCCATCGATACGACCCGCTA	1219	2240	GCCCTTGTGATGGAGGGAAGAACTCTACTGTGGGCGGACACTGGGACCAACAGATCGAA	2299
Qy	1201	GAGGCTATGTCTACTGACAGATGACGAGTGGGGCCATCCGAGGGCGTACTGGAC	1260	2281	GTGGCGGGTGGACGGGCGAGTCCGCGCAAGTCCCTGTGTGGAGGAGCTTGGACCAACCG	2340
Db	1220	GAGGCTATGTCTACTGACAGATGACGAGTGGGGCCATCCGAGGGCGTACTGGAC	1279	2300	GTGGCGGGTGGACGGGCGAGTCCGCGCAAGTCCCTGTGTGGAGGAGCTTGGACCAACCG	2359
Qy	1261	GGGTCTGGGGCGAGAGCTGTCAACACCGAGATCAACGACCCCGATGGCATCGCGTC	1320	2341	AGGTCTGTGGCGCTTGGATCCCAACAGGGCTTACATCTACTGGAACGAGTGGGGCGGCAAG	2400
Db	1280	GGGTCTGGGGCGAGAGCTGTCAACACCGAGATCAACGACCCCGATGGCATCGCGTC	1339	2360	AGGTCTGTGGCGCTTGGATCCCAACAGGGCTTACATCTACTGGAACGAGTGGGGCGGCAAG	2419
Qy	1321	GACTGGGTGGCCGAAACCTCTACTGGACCGACCGGACCGGACCGCATCGAGTGAAG	1380	2401	CCGAGGATCTGTGGGGCTTCAATGGAACGGGACCAACTGATGACGCTGTGGTGGACAGGTG	2460
Db	1340	GACTGGGTGGCCGAAACCTCTACTGGACCGACCGGACCGGACCGCATCGAGTGAAG	1399	2420	CCGAGGATCTGTGGGGCTTCAATGGAACGGGACCAACTGATGACGCTGTGGTGGACAGGTG	2479
Qy	1381	CGCTCAACGACACTCCCGCAAGATCTGTGTGCGAGGACCTGAGAGCCCGGAGCC	1440	2461	GGCGGGCGGACGACCTCAACATTTGACTAGCTGACGAGCGGCTTACTGACCGGACCTG	2520
Db	1400	CGCTCAACGACACTCCCGCAAGATCTGTGTGCGAGGACCTGAGAGCCCGGAGCC	1459	2480	GGCGGGCGGACGACCTCAACATTTGACTAGCTGACGAGCGGCTTACTGACCGGACCTG	2539
Qy	1441	ATCGCACTGACCCCGTATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAA	1500	2521	GACACCAACATGATCGAGTGTCCCAACATGCTGGGTGAGGAGCGGCTGTGATTTGCCAC	2580
Db	1460	ATCGCACTGACCCCGTATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAA	1519	2540	GACACCAACATGATCGAGTGTCCCAACATGCTGGGTGAGGAGCGGCTGTGATTTGCCAC	2599
Qy	1501	ATCGAGTGTGCCAACTTGGATGGGACGAGCGGCTGTGTGCTCAATGCTTCCCTCGG	1560	2581	GATCTCCGCAACCCGCTTGGTGTGAGCGAGTACAGCGATTTATCTACTGACAGACTGG	2640
Db	1520	ATCGAGTGTGCCAACTTGGATGGGACGAGCGGCTGTGTGCTCAATGCTTCCCTCGG	1579	2600	GATCTCCGCAACCCGCTTGGTGTGAGCGAGTACAGCGATTTATCTACTGACAGACTGG	2659
Qy	1561	TGGCCCAACGCGCTGGCCCTGGAACCTGACAGAGGGAGCTTACTTGGGAGAGCGCCAG	1620	2641	AATCTGCACAGCATTTGAGCGGGCGGACAAAGTACTAGCGGCGGAAACCGACCTTATCCAG	2700
Db	1580	TGGCCCAACGCGCGCTGGCCCTGGAACCTGACAGAGGGAGCTTACTTGGGAGAGCGCCAG	1639	2660	AATCTGCACAGCATTTGAGCGGGCGGACAAAGTACTAGCGGCGGAAACCGACCTTATCCAG	2719
Qy	1621	ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCTCTCGGAGGAC	1680	2701	GGCCACCTGGACTTCTGTGATGGACATCTGTGTGTTTCCACTCTCTCCGCGAGGATGGCTC	2760
Db	1640	ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCTCTCTCGGAGGAC	1699	2720	GGCCACCTGGACTTCTGTGATGGACATCTGTGTGTTTCCACTCTCTCCGCGAGGATGGCTC	2779
Qy	1681	AAGCTCCGCGCAATTTTCGGGTTCACGCTGTGGGGGACTTCACTACTGAGTGAATGG	1740	2761	AATGACTGTATGCAACAAACGGGCGAGTGTGGGAGCTGTGCTTGCATCCCGCGGCGC	2820
Db	1700	AAGCTCCGCGCAATTTTCGGGTTCACGCTGTGGGGGACTTCACTACTGAGTGAATGG	1759	2780	AATGACTGTATGCAACAAACGGGCGAGTGTGGGAGCTGTGCTTGCATCCCGCGGCGC	2839
Qy	1741	CAGCGCGCAGCATCGAGGGTGCACAGGTCAAGGCGCGGAGCTCATATTGAC	1800	2821	CACCGCTGGGCTCGGCTCACTACACTTACCTTGGACCCAGCAGCGGCACTGACGCGCG	2880
Db	1760	CAGCGCGCAGCATCGAGGGTGCACAGGTCAAGGCGCGGAGCTCATATTGAC	1819	2840	CACCGCTGGGCTCGGCTCACTACACTTACCTTGGACCCAGCAGCGGCACTGACGCGCG	2899
				2881	CCCACCACTTCTTGTGCTTTCAGCCAGAAATCTGCCATCTAGTCCGATGATCCCGGACGAC	2940

Qy	1681	AAGCTCCCGCACATTTTCGGGGTTCACGCTGCTGGGGGACTTCACTACTGACTGACTGG	1740
Db	1700	AAGCTCCCGCACATTTTCGGGGTTCACGCTGCTGGGGGACTTCACTACTGACTGACTGG	1759
Qy	1741	CAGCGCGCAGCATCGAGCGGGTGCAAGGTCAAGGCTCAAGGCGAGCGGAGCGTCACTATTGAC	1800
Db	1760	CAGCGCGCAGCATCGAGCGGGTGCAAGGTCAAGGCTCAAGGCGAGCGGAGCGTCACTATTGAC	1819
Qy	1801	CAGCTGCCCGACCATGATGGGGCTCAAGCTGTGAATGTGGCCAAGGTGCTGGAAACCAAC	1860
Db	1820	CAGCTGCCCGACCATGATGGGGCTCAAGCTGTGAATGTGGCCAAGGTGCTGGAAACCAAC	1879
Qy	1861	CCGTGTGGGACAGGAACCGGGGGTGACACCACTGTCTTTTTCACACCCCAAGCAACC	1920
Db	1880	CCGTGTGGGACAGGAACCGGGGGTGACACCACTGTCTTTTTCACACCCCAAGCAACC	1939
Qy	1921	CGGTGTGGCTGCCCATCGGCTCGAGCTGTGAGTGATGAAGACCTGCACTGTGCCT	1980
Db	1940	CGGTGTGGCTGCCCATCGGCTCGAGCTGTGAGTGATGAAGACCTGCACTGTGCCT	1999
Qy	1981	GAGGCTTCTTTGGTCTTTCACAGCAGAGCGCGCCATCCACAGGATCTCCCTCGAGACCAAT	2040
Db	2000	GAGGCTTCTTTGGTCTTTCACAGCAGAGCGCGCCATCCACAGGATCTCCCTCGAGACCAAT	2059
Qy	2041	AAACACGACGTGGCCATCCCGCTCAGCGGGCGTCAAGGAGGCGCTCAGCGCCTCGACTTTTGAT	2100
Db	2060	AAACACGACGTGGCCATCCCGCTCAGCGGGCGTCAAGGAGGCGCTCAGCGCCTCGACTTTTGAT	2119
Qy	2101	GTGTCCAAACCAACATCTACTGGAAGACGCTGAGCTGGAAGACCATCAGCGCGGCTTC	2160
Db	2120	GTGTCCAAACCAACATCTACTGGAAGACGCTGAGCTGGAAGACCATCAGCGCGGCTTC	2179
Qy	2161	ATGAACGGGAGCTCGGTGGAGCACTGTGTGAGTTTGGCTTTGACTACCTCCCGAGGGCATG	2220
Db	2180	ATGAACGGGAGCTCGGTGGAGCACTGTGTGAGTTTGGCTTTGACTACCTCCCGAGGGCATG	2239
Qy	2221	GCCGTTGACTGAGTGGGCAAGAACTCTACTTGGCGGACACTGGGACCAACAGAAATCGAA	2280
Db	2240	GCCGTTGACTGAGTGGGCAAGAACTCTACTTGGCGGACACTGGGACCAACAGAAATCGAA	2299
Qy	2281	GTGCGCGGCTGGACGGGCAAGTCCGGCAAGTCTCTGTGTGGAGGGACTTGGACAAACCGG	2340
Db	2300	GTGCGCGGCTGGACGGGCAAGTCCGGCAAGTCTCTGTGTGGAGGGACTTGGACAAACCGG	2359
Qy	2341	AGTCTGCTGGCCCTGGATCCCAACAGGGCTACATCTACTTGGACCGAGTGGGGCGGCAAG	2400
Db	2360	AGTCTGCTGGCCCTGGATCCCAACAGGGCTACATCTACTTGGACCGAGTGGGGCGGCAAG	2419
Qy	2401	CCGAGGATCGTGGCGGCTTTCATGGACGGGACCAACTGATGACGCTGGTGGAACAAGTG	2460
Db	2420	CCGAGGATCGTGGCGGCTTTCATGGACGGGACCAACTGATGACGCTGGTGGAACAAGTG	2479
Qy	2461	GGCGGGCCAAACGACTCAACATTGACTACGCTGACCAAGCGCTCTACTTGGACCGACTG	2520
Db	2480	GGCGGGCCAAACGACTCAACATTGACTACGCTGACCAAGCGCTCTACTTGGACCGACTG	2539
Qy	2521	GACACCAACATGATCGAGTGTCTCAACATGTCTGGGTGAGGACGGGTGCTGATGGCCAC	2580
Db	2540	GACACCAACATGATCGAGTGTCTCAACATGTCTGGGTGAGGACGGGTGCTGATGGCCAC	2599
Qy	2581	GATCTCCCGCACCGCTTGGGTCTGACGCAAGTACAGCGATTATATCTACTTGGACAGACTGG	2640
Db	2600	GATCTCCCGCACCGCTTGGGTCTGACGCAAGTACAGCGATTATATCTACTTGGACAGACTGG	2659
Qy	2641	AATCTGACAGCATTTGAGCGGGCCGACAAAGACTAGCGCGCGGAACCGCAACCTCATCCAG	2700
Db	2660	AATCTGACAGCATTTGAGCGGGCCGACAAAGACTAGCGCGCGGAACCGCAACCTCATCCAG	2719
Qy	2701	GGCCACCTTGGACTGTGATGGACATCTGTGTGTTCATCTCTCCCGCAGGATGGGCTC	2760
Db	2720	GGCCACCTTGGACTGTGATGGACATCTGTGTGTTCATCTCTCCCGCAGGATGGGCTC	2779

Qy	2761	AATGACTGTATGACAAACAAACGGGCAAGTGTGGGAGCTGTGCCCTTGGCCATCCCGCGGCGGC	2820
Db	2780	AATGACTGTATGACAAACAAACGGGCAAGTGTGGGAGCTGTGCCCTTGGCCATCCCGCGGCGGC	2839
Qy	2821	CACCGCTGGGCTGGCTCTCACTACACCTTGGACCCAGCAGCCGCAACTGCAAGCCCG	2880
Db	2840	CACCGCTGGGCTGGCTCTCACTACACCTTGGACCCAGCAGCCGCAACTGCAAGCCCG	2899
Qy	2881	CCGACCACTTCTTGTCTTTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGGACGAC	2940
Db	2900	CCGACCACTTCTTGTCTTTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGGACGAC	2959
Qy	2941	CAGCAGCGCCCGGACTCTCATCTGCTCCCTGATGAGACTGAGAAAGCTCAAGACCACTCCAG	3000
Db	2960	CAGCAGCGCCCGGACTCTCATCTGCTCCCTGATGAGACTGAGAAAGCTCAAGACCACTCCAG	3019
Qy	3001	TATGACCACTTGGACAAAGTTTCATCTACTGGTGTGATGGCGCCAGAAACATCAAGCGAGCC	3060
Db	3020	TATGACCACTTGGACAAAGTTTCATCTACTGGTGTGATGGCGCCAGAAACATCAAGCGAGCC	3079
Qy	3061	AAGGACGAGCGGACCCAGCCCTTTTGTGTGACTCTCTGAGCCAGGCAAAACCCAGAC	3120
Db	3080	AAGGACGAGCGGACCCAGCCCTTTTGTGTGACTCTCTGAGCCAGGCAAAACCCAGAC	3139
Qy	3121	AGGAGCGCCCAACGACCTCAGCATCGACATCTACAGCCGCGACACTGTCTTGGACGTGCCAG	3180
Db	3140	AGGAGCGCCCAACGACCTCAGCATCGACATCTACAGCCGCGACACTGTCTTGGACGTGCCAG	3199
Qy	3181	GCCACCAATACCATCAACGCTCCACAGCTGAGCGGGGAGCCATGGGGGTGGTGTGGT	3240
Db	3200	GCCACCAATACCATCAACGCTCCACAGCTGAGCGGGGAGCCATGGGGGTGGTGTGGT	3259
Qy	3241	GGGACCGCGACAGCCGAGCCGAGCCATCGTGTCTAAACGGGAGCGAGGCTACTGTACTTC	3300
Db	3260	GGGACCGCGACAGCCGAGCCGAGCCATCGTGTCTAAACGGGAGCGAGGCTACTGTACTTC	3319
Qy	3301	ACCAACATGAGGACCGGGGAGCCAAAGTCCAAACGCGAGCCCTTGGACGCAACGAGCGC	3360
Db	3320	ACCAACATGAGGACCGGGGAGCCAAAGTCCAAACGCGAGCCCTTGGACGCAACGAGCGC	3379
Qy	3361	GAGTCTCTTTCACCAACCGGCTCATCGGCTGTGCGCTGTGGTGTGGACAAACACATG	3420
Db	3380	GAGTCTCTTTCACCAACCGGCTCATCGGCTGTGCGCTGTGGTGTGGACAAACACATG	3439
Qy	3421	GGCAAGCTGTCTGGGTGGAGCGGACCTGAAAGCGCATTTGAGCTGTGACTGTCTGAGG	3480
Db	3440	GGCAAGCTGTCTGGGTGGAGCGGACCTGAAAGCGCATTTGAGCTGTGACTGTCTGAGG	3499
Qy	3481	GCCAAACCGCTGACCCCTGGAGGACGCCAAACATCGTGGAGCCCTCTGGGCTTGAACCATCTT	3540
Db	3500	GCCAAACCGCTGACCCCTGGAGGACGCCAAACATCGTGGAGCCCTCTGGGCTTGAACCATCTT	3559
Qy	3541	GGCAAGCATCTTACTGGATCGACCGCAGCAGCAGATGATCGAGCGTGTGGAGAGACC	3600
Db	3560	GGCAAGCATCTTACTGGATCGACCGCAGCAGCAGATGATCGAGCGTGTGGAGAGACC	3619
Qy	3601	ACCGGGGACAGCGGACTCGCATCCAGGCGCTGTGCGCCACCTCACTGGGATCCATGCA	3660
Db	3620	ACCGGGGACAGCGGACTCGCATCCAGGCGCTGTGCGCCACCTCACTGGGATCCATGCA	3679
Qy	3661	GTGAGGAAAGTTCAGCCTGGAGGAGTTCTTACGCCCAACCATGTGCGCGGTGACAAATGGTGGC	3720
Db	3680	GTGAGGAAAGTTCAGCCTGGAGGAGTTCTTACGCCCAACCATGTGCGCGGTGACAAATGGTGGC	3739
Qy	3721	TGCTCCCAACATCTGTATTGCAAGGGTGTGGGACACCAACGCTGTCTATGCCCAAGTCCAC	3780
Db	3740	TGCTCCCAACATCTGTATTGCAAGGGTGTGGGACACCAACGCTGTCTATGCCCAAGTCCAC	3799
Qy	3781	CTCGTGTCTTCGCAAGACCTGCTGACCTGTGGAGCGCGCCACCTGCTCTCCCGGACGAC	3840
Db	3800	CTCGTGTCTTCGCAAGACCTGCTGACCTGTGGAGCGCGCCACCTGCTCTCCCGGACGAC	3859
Qy	3841	TTTGCATGTGCAACAGGGAGATCGACTGTATCCCGGGGCTTGGCGCTGTGTGACGGCTTT	3900

Db	3860	TTTGCATGTGCCACAGGGGAGATCGACTGTATATCCCGGGGCGCTCGCGCTGTGACGGCTTT	3919
Qy	3901	CCGAGTGCATGATCACAGAGCGACGAGGAGGCGCTGCCCGGTGTGCTCCGCGCCCGCAGTTC	3960
Db	3920	CCCGAGTGCATGATCACAGAGCGACGAGGAGGCGCTGCCCGGTGTGCTCCGCGCCCGCAGTTC	3979
Qy	3961	CCCTGCGGCGGGGTCAGTGTGTGGAAGCTTGGCCCTGTGCTGCGACGCGGAGGCGAGACTGT	4020
Db	3980	CCCTGCGGCGGGGTCAGTGTGTGGAAGCTTGGCCCTGTGCTGCGACGCGGAGGCGAGACTGT	4039
Qy	4021	CAGGACCGCTCAGACGAGGCGGAGCTGTGTGACGCCCATCTGCTGCCCAACAGTTCGCGTGT	4080
Db	4040	CAGGACCGCTCAGACGAGGCTGTGTGTGACGCCCATCTGCTGCCCAACAGTTCGCGTGT	4099
Qy	4081	CGGAGCGGCTCAGTGTGTCTCATCAAAACAGAGTGTGCACTCTTCCCGGACTGTATCGAC	4140
Db	4100	CGGAGCGGCTCAGTGTGTCTCATCAAAACAGAGTGTGCACTCTTCCCGGACTGTATCGAC	4159
Qy	4141	GGCTCCGACGAGCTCATGTGTGAAATCACCAAGCGCGCTCAGACGACGCCGCGGCCAC	4200
Db	4160	GGCTCCGACGAGCTCATGTGTGAAATCACCAAGCGCGCTCAGACGACGCCGCGGCCAC	4219
Qy	4201	AGCAGTGCATCGGCGCCCGTCAATTGGCATCATCTCTCTCTTTCGTCAATGGGTGTGTC	4260
Db	4220	AGCAGTGCATCGGCGCCCGTCAATTGGCATCATCTCTCTCTTTCGTCAATGGGTGTGTC	4279
Qy	4261	TATTTTGTGTGCCAGCGCGTGTGTGCCAGGCTATGCGGGGGCGCAACGGGCGCTTCCCG	4320
Db	4280	TATTTTGTGTGCCAGCGCGTGTGTGCCAGGCTATGCGGGGGCGCAACGGGCGCTTCCCG	4339
Qy	4321	CACGAGTATGTACGCGGGAACCCGCAAGTGCCTCTCAATTTTATAGCCCCGGGCGGTTC	4380
Db	4340	CACGAGTATGTACGCGGGAACCCGCAAGTGCCTCTCAATTTTATAGCCCCGGGCGGTTC	4399
Qy	4381	CAGCATGGCCCTTTCACAGGATGTGATGCGGAAGTTCATGATGAGTCCGTGAGCCTG	4440
Db	4400	CAGCATGGCCCTTTCACAGGATGTGATGCGGAAGTTCATGATGAGTCCGTGAGCCTG	4459
Qy	4441	ATGGGGGCGGGGCGGGTGCCTCTGTACGCGGGAACCACTGTACAGGGGCGCTCGTCC	4500
Db	4460	ATGGGGGCGGGGCGGGTGCCTCTGTACGCGGGAACCACTGTACAGGGGCGCTCGTCC	4519
Qy	4501	AGCAGTCTGTCCAGACGAAGGCGACGCTGTACCGCGGATCTGTGAACCGCGCGCTCC	4560
Db	4520	AGCAGTCTGTCCAGACGAAGGCGACGCTGTACCGCGGATCTGTGAACCGCGCGCTCC	4579
Qy	4561	CCGSCCAGCAGCCCTCCCTGTACACATGCAATGTCTTACTCTTTCACATTCGCGGC	4620
Db	4580	CCGSCCAGCAGCCCTCCCTGTACACATGCAATGTCTTACTCTTTCACATTCGCGGC	4639
Qy	4621	ACTGCGAGCCGTACAGGCGCTACATCATTCGAGGAATGGCGCCCCCGACGAGCGCGCTGC	4680
Db	4640	ACTGCGAGCCGTACAGGCGCTACATCATTCGAGGAATGGCGCCCCCGACGAGCGCGCTGC	4699
Qy	4681	AGCACCGAGTGTGTGACGCGACTACAGCGCCAGCGCTGTGGAAGGCGGACGAGTACTAC	4740
Db	4700	AGCACCGAGTGTGTGACGCGACTACAGCGCCAGCGCTGTGGAAGGCGGACGAGTACTAC	4759
Qy	4741	CTGGAATTGAATCTCGGACTCAGACCCCTATTCACCCCGACCCCGGCGGCGCGCGCTGC	4800
Db	4760	CTGGAATTGAATCTCGGACTCAGACCCCTATTCACCCCGACCCCGGCGGCGCGCGCTGC	4819
Qy	4801	CTGTGCGCGAGGACAGCTGCGCGCTCGCGCCACCGAGAGGAGTACTTTCATCTC	4860
Db	4820	CTGTGCGCGAGGACAGCTGCGCGCGCTCGCGCCACCGAGAGGAGTACTTTCATCTC	4879
Qy	4861	TTCCGCGCCCTCCGTCCTCCAGCGACTCATCTGACCTCGCGCGGCGGCGACTCTGCG	4920
Db	4880	TTCCGCGCCCTCCGTCCTCCAGCGACTCATCTGACCTCGCGCGGCGGCGACTCTGCG	4939
Qy	4921	TTCTCTGTGCGCCCTGTAAATAGTTTTAAATATGAAACAAAGAAAAAATATATTTATGAT	4980

[illegible]

Qy	421	TGGACGAGCTCAGAGACCAACCGCATCGAGGTGGCCAACTCAATGGCACATCCCGAAG	480	Qy	1501	ATCGAGTGTGCAACTTTGGATGGGACGAGCGGCTGTGCTGGTCAATGCCCTCCCTCGGG	1560
Db	440	TGGACGAGCTCAGAGACCAACCGCATCGAGGTGGCCAACTCAATGGCACATCCCGAAG	499	Db	1520	ATCGAGTGTGCAACTTTGGATGGGACGAGCGGCTGTGCTGGTCAATGCCCTCCCTCGGG	1579
Qy	481	GTGCTCTTTCTGGCAGGACCTTGAACAGCTGAGGCGCATCGCTTGGACCCCGCTCACGGG	540	Qy	1561	TGGCCCAACGGCCTTGGCCCTGGACCTTGACAGAGGGGAAGCTTACTTGGGGAGACCCCAAG	1620
Db	500	GTGCTCTTTCTGGCAGGACCTTGAACAGCTGAGGCGCATCGCTTGGACCCCGCTCACGGG	559	Db	1580	TGGCCCAACGGCCTTGGCCCTGGACCTTGACAGAGGGGAAGCTTACTTGGGGAGACCCCAAG	1639
Qy	541	TACATGTACTGACAGACTGGGGTGGAGCGCCCGGATTGAGCGGCGAGGATGATGGC	600	Qy	1621	ACAGCAAGATCGAGGTGATCAATGTTGATCGGACGAAGAGCGGACCTCTCTGGAGGAC	1680
Db	560	TACATGTACTGACAGACTGGGGTGGAGCGCCCGGATTGAGCGGCGAGGATGATGGC	619	Db	1640	ACAGCAAGATCGAGGTGATCAATGTTGATCGGACGAAGAGCGGACCTCTCTGGAGGAC	1699
Qy	601	AGCACCCGGAAGATCAATTTGTGACTCGGACATTTTACTTGGCCCAATGGACTGACCATCGAC	660	Qy	1681	AAGCTCCCGCACATTTTTCGGGTTTCAACGCTGTGGGGGACTTTCATCTACTGAGTGAATGG	1740
Db	620	AGCACCCGGAAGATCAATTTGTGACTCGGACATTTTACTTGGCCCAATGGACTGACCATCGAC	679	Db	1700	AAGCTCCCGCACATTTTTCGGGTTTCAACGCTGTGGGGGACTTTCATCTACTGAGTGAATGG	1759
Qy	661	CTGAGGAGCAGAGCTCTACTTGTGACTGAGCAGACTGGCAGACCCGCTCCATCCATGCGCTGC	720	Qy	1741	CAGCGCCGACGATCGAGCGGGTGCACAGGTCTAAGGCTCAAGGCGAGCGGGACGTCACTATTGAC	1800
Db	680	CTGAGGAGCAGAGCTCTACTTGTGACTGAGCAGACTGGCAGACCCGCTCCATCCATGCGCTGC	739	Db	1760	CAGCGCCGACGATCGAGCGGGTGCACAGGTCTAAGGCTCAAGGCGAGCGGGACGTCACTATTGAC	1819
Qy	721	CTGAGCGCTCTTCCGCGAGAGGTGTGGAGGCGAGCTGACGCAACCCCTTGGCCCTG	780	Qy	1801	CAGCTTGCCTGACCTGATGGGCTCAAAGCTGTGTAATGTGGCAAGGCTCGTGGAAACCAAC	1860
Db	740	CTGAGCGCTCTTCCGCGAGAGGTGTGGAGGCGAGCTGACGCAACCCCTTGGCCCTG	799	Db	1820	CAGCTTGCCTGACCTGATGGGCTCAAAGCTGTGTAATGTGGCAAGGCTCGTGGAAACCAAC	1879
Qy	781	ACGCTCTCCGGGGACACTCTGTACTTGGACAGACTGGCAGACCCGCTCCATCCATGCGCTGC	840	Qy	1861	CCGTGTGCGGACAGAAACGGGGGTGCAGCCACCTGTGCTTTTTCACACCCCAACCAACC	1920
Db	800	ACGCTCTCCGGGGACACTCTGTACTTGGACAGACTGGCAGACCCGCTCCATCCATGCGCTGC	859	Db	1880	CCGTGTGCGGACAGAAACGGGGGTGCAGCCACCTGTGCTTTTTCACACCCCAACCAACC	1939
Qy	841	AACAAAGCGCACTGGGGGAGAGGAGAGATCTCTGAGTGCCTCTACTCAACCATGGAC	900	Qy	1921	CGGTGTGGCTGCCCATCGGCTTGGAGCTGTGATGATGACATGAAGACCTTGATCGTGCCT	1980
Db	860	AACAAAGCGCACTGGGGGAGAGGAGAGATCTCTGAGTGCCTCTACTCAACCATGGAC	919	Db	1940	CGGTGTGGCTGCCCATCGGCTTGGAGCTGTGATGATGACATGAAGACCTTGATCGTGCCT	1999
Qy	901	ATCAGGTGCTGAGCCAGGCGGAGCTTTCTTCCACACTCGCTGTGAGGAGACAAAT	960	Qy	1981	GAGGCTTTCTTGTGCTTCCACGACAGAGCGCCCATCCACAGATCTCCCTCGAGACCAAT	2040
Db	920	ATCAGGTGCTGAGCCAGGCGGAGCTTTCTTCCACACTCGCTGTGAGGAGACAAAT	979	Db	2000	GAGGCTTTCTTGTGCTTCCACGACAGAGCGCCCATCCACAGATCTCCCTCGAGACCAAT	2059
Qy	961	GGCGCTGCTCCCACTGTGCTGTGCTCCCAAGCGAGCGCTTTCTACACATGGCGCTGC	1020	Qy	2041	AACAAAGCGCACTGGGGCATCCCGCTCAACGGGCGCTCAAGGAGGCTCAGCCCTGAGCTTTGAT	2100
Db	980	GGCGCTGCTCCCACTGTGCTGTGCTCCCAAGCGAGCGCTTTCTACACATGGCGCTGC	1039	Db	2060	AACAAAGCGCTGGCCATCCCGCTCAACGGGCGCTCAAGGAGGCTCAGCCCTGAGCTTTGAT	2119
Qy	1021	CCACGGGTGTGAGCTGACAGCAAAACGGCAGGAGCTGTAAAGCAGGAGCGGAGAGGTG	1080	Qy	2101	GTGTCCCAACCAACCATCTACTTGACAGACGTCTAGCTCTGAAGACCATCAGCGCGGCTTC	2160
Db	1040	CCACGGGTGTGAGCTGACAGCAAAACGGCAGGAGCTGTAAAGCAGGAGCGGAGAGGTG	1099	Db	2120	GTGTCCCAACCAACCATCTACTTGACAGACGTCTAGCTCTGAAGACCATCAGCGCGGCTTC	2179
Qy	1081	CTGCTGCTGGCCCGCGGACCGACCTACGAGGATCTCGCTGGACACGCCCGGACTTCACC	1140	Qy	2161	ATGAACGGGAGCTCGGTGGAGCAGTGTGAGTTTGGCCCTTGAATACCCCGAGGCGATG	2220
Db	1100	CTGCTGCTGGCCCGCGGACCGACCTACGAGGATCTCGCTGGACACGCCCGGACTTCACC	1159	Db	2180	ATGAACGGGAGCTCGGTGGAGCAGTGTGAGTTTGGCCCTTGAATACCCCGAGGCGATG	2239
Qy	1141	GACATCGTGTGACAGGTGACAGACATCCCGCAGCGCATTTGCCATCGACTACGACCCGCTA	1200	Qy	2221	GCCGTTGACTGGATGGGCAAGAACTCTACTTGGGCGGACACTGGGACCAACAGATCGAA	2280
Db	1160	GACATCGTGTGACAGGTGACAGACATCCCGCAGCGCATTTGCCATCGACTACGACCCGCTA	1219	Db	2240	GCCGTTGACTGGATGGGCAAGAACTCTACTTGGGCGGACACTGGGACCAACAGATCGAA	2299
Qy	1201	GAGGCTATGTCTACTGACAGATGACAGGTGCGGGCCATCCGACAGGCGCTACTGGAC	1260	Qy	2281	GTGGCGGCTGAGACGGGCAAGTTCCGGCAAGTCTCTGTGTGGAGGGAATTGGACCAACCGG	2340
Db	1220	GAGGCTATGTCTACTGACAGATGACAGGTGCGGGCCATCCGACAGGCGCTACTGGAC	1279	Db	2300	GTGGCGGCTGAGACGGGCAAGTTCCGGCAAGTCTCTGTGTGGAGGGAATTGGACCAACCGG	2359
Qy	1261	GGGTGTGGGGCGAGACGCTGTGTAACACCGAGATCAACGACCCCGATGGCATCGCGTC	1320	Qy	2341	AGGTGCTTGGCCCTTGGATCCCAACAGGCTTACTCTGAGCCGAGTGGGGCGGCAAG	2400
Db	1280	GGGTGTGGGGCGAGACGCTGTGTAACACCGAGATCAACGACCCCGATGGCATCGCGTC	1339	Db	2360	AGGTGCTTGGCCCTTGGATCCCAACAGGCTTACTCTGAGCCGAGTGGGGCGGCAAG	2419
Qy	1321	GACTGGGTGGCCCGAAACCTCTACTTGGACCGACACGGGACCGGACCGCATCGAGGTGACG	1380	Qy	2401	CCGAGGATCGTGGCGGCTTTCATGGACGGGACCAACTGATGACGCTGTGTGGAGGAGGTTG	2460
Db	1340	GACTGGGTGGCCCGAAACCTCTACTTGGACCGACACGGGACCGGACCGCATCGAGGTGACG	1399	Db	2420	CCGAGGATCGTGGCGGCTTTCATGGACGGGACCAACTGATGACGCTGTGTGGAGGAGGTTG	2479
Qy	1381	CGCTCAACGGACCTCCCGCAAGATCTGTGTGAGGAGCTTGGACAGACCCCGAGCC	1440	Qy	2461	GGCCGGGCAACGACCTTCAACATTTGACTACGCTGACGCGCTCTACTTGACCGACCTG	2520
Db	1400	CGCTCAACGGACCTCCCGCAAGATCTGTGTGAGGAGCTTGGACAGACCCCGAGCC	1459	Db	2480	GGCCGGGCAACGACCTTCAACATTTGACTACGCTGACGCGCTCTACTTGACCGACCTG	2539
Qy	1441	ATCGACTGCAACCCGCTGATGGGCTCTATGTACTTGGACAGACTGGGAGAGAACCTTAA	1500	Qy	2521	GACACCAACATGATCGAGTCTGTCACATGCTGGGTCAAGGAGCGGCTGTGATTCGCCAC	2580
Db	1460	ATCGACTGCAACCCGCTGATGGGCTCTATGTACTTGGACAGACTGGGAGAGAACCTTAA	1519	Db	2540	GACACCAACATGATCGAGTCTGTCACATGCTGGGTCAAGGAGCGGCTGTGATTCGCCAC	2599
				Qy	2581	GATCTCCGCAACCCGCTTCCGCTCTGACGCACTACAGCGATTATCTACTGAGACAGACTGG	2640

Db 2600 GATCTCCGCAACCCGTTGGTCTGACGCGAGTACAGGATATATCTACTGGACAGACTGG 2659
QY 2641 AATCTGCACAGCAATGAGCGGCGGACAGACTAGCGCGGAAACCGACCCCTCATCCAG 2700
Db 2660 AATCTGCACAGCAATGAGCGGCGGACAGACTAGCGCGGAAACCGACCCCTCATCCAG 2719
QY 2701 GGCACCTGGACTTGGTATGACATCTCTGGTGTTCACATCTCTCCGGCAGGATGGCCTC 2760
Db 2720 GGCACCTGGACTTGGTATGACATCTCTGGTGTTCACATCTCTCCGGCAGGATGGCCTC 2779
QY 2761 AATGACTGATGACAAACAAACGGGAGTGTGGGAGCTGTGCTTGGCCATCCCGCGCGAC 2820
Db 2780 AATGACTGATGACAAACAAACGGGAGTGTGGGAGCTGTGCTTGGCCATCCCGCGCGAC 2839
QY 2821 CACCGCTGGGTGGCTCTACATCTACACCTCTGACAGCTGAGGAACTCTCAAGGCCATCGAC 2880
Db 2840 CACCGCTGGGTGGCTCTACATCTACACCTCTGACAGCTGAGGAACTCTCAAGGCCATCGAC 2899
QY 2881 CCACACCTCTCTGCTGTTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGGACGAC 2940
Db 2900 CCACACCTCTCTGCTGTTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGGACGAC 2959
QY 2941 CAGCACAGCCCGGATCTCATCTCTGCCCTCTGATGAGCTGAGGAACTCTCAAGGCCATCGAC 3000
Db 2960 CAGCACAGCCCGGATCTCATCTCTGCCCTCTGATGAGCTGAGGAACTCTCAAGGCCATCGAC 3019
QY 3001 TATGACCACTGGAACAAGTTTCACTACTGGGTGGATGGCGGCAGAAATCAAGCGAGCC 3060
Db 3020 TATGACCACTGGAACAAGTTTCACTACTGGGTGGATGGCGGCAGAAATCAAGCGAGCC 3079
QY 3061 AAGGACGAGGAGCCAGCCCTTTGTTTGAACCTCTCTGAGCCAGGCGCAACCCAGAC 3120
Db 3080 AAGGACGAGGAGCCAGCCCTTTGTTTGAACCTCTCTGAGCCAGGCGCAACCCAGAC 3139
QY 3121 AAGGACGAGGAGCCAGCCCTCTGATCGACATCTACAGCCGGAACACTTCTTGGACGTGGCGAG 3180
Db 3140 AAGGACGAGGAGCCAGCCCTCTGATCGACATCTACAGCCGGAACACTTCTTGGACGTGGCGAG 3199
QY 3181 GCCACCAATACATCAACGTTCACAGGCTGAGCGGGGAAGCATGGGGGTGGTGTGGT 3240
Db 3200 GCCACCAATACATCAACGTTCACAGGCTGAGCGGGGAAGCATGGGGGTGGTGTGGT 3259
QY 3241 GGGGACCGGACAGCCAGGSCATCTGCTCAACGGGAGCGAGGCTACTGTACTTTC 3300
Db 3260 GGGGACCGGACAGCCAGGSCATCTGCTCAACGGGAGCGAGGCTACTGTACTTTC 3319
QY 3301 ACCAACATGCAAGGACCGGGCAGCCAAAGATCGAAACGGCAGCCCTGGACCGGACGCGC 3360
Db 3320 ACCAACATGCAAGGACCGGGCAGCCAAAGATCGAAACGGCAGCCCTGGACCGGACGCGC 3379
QY 3361 GAGGTCTCTTTCACACCGGCTCATCGCCCTGTGGCCCTGTGGTGGGACAAACACTG 3420
Db 3380 GAGGTCTCTTTCACACCGGCTCATCGCCCTGTGGCCCTGTGGTGGGACAAACACTG 3439
QY 3421 GCGAAGCTGTTCTGGGTGGACGCGACCTGAAGCGCATTTGAGAGCTGTGACTGTGAGG 3480
Db 3440 GCGAAGCTGTTCTGGGTGGACGCGACCTGAAGCGCATTTGAGAGCTGTGACTGTGAGG 3499
QY 3481 GCCAACCGCCTGACCTTGGAGGACGCCAACTCGTGCAGCCTCTGGGCTTGACCAATCCTT 3540
Db 3500 GCCAACCGCCTGACCTTGGAGGACGCCAACTCGTGCAGCCTCTGGGCTTGACCAATCCTT 3559
QY 3541 GCGAGGATCTCTACTGATGACCGCCAGCAGCAGATGATCGAGCTGTGGAGAGACC 3600
Db 3560 GCGAGGATCTCTACTGATGACCGCCAGCAGCAGATGATCGAGCTGTGGAGAGACC 3619
QY 3601 ACCGGGGAACAGCGGACTCGATTCAGGGCCGTGTGCGCCACCTCACTGGGATCCATGCA 3660
Db 3620 ACCGGGGAACAGCGGACTCGATTCAGGGCCGTGTGCGCCACCTCACTGGGATCCATGCA 3679
QY 3661 GTGGAGGAAGTACGCTCGAGGAGTTCTCAGCCCAACCATGTGCCCGTGAACAATGGTGGC 3720

Db 3680 GTGGAGGAAGTACGCTCGAGGAGTTCTCAGCCCAACCATGTGCCCGTGACAAATGGTGGC 3739
QY 3721 TGCTCCCAATCTGTATTTGCCAAGGTGATGGGACACACAGGTGTCTATGGCCAGTCCAC 3780
Db 3740 TGCTCCCAATCTGTATTTGCCAAGGTGATGGGACACACAGGTGTCTATGGCCAGTCCAC 3799
QY 3781 CTGCTGCTCTCCGACAACTCTGACCTGTGTGAGAGCGCCCACTGTCTCCCGGACGAC 3840
Db 3800 CTGCTGCTCTCCGACAACTCTGACCTGTGTGAGAGCGCCCACTGTCTCCCGGACGAC 3859
QY 3841 TTTGATGTGCAACAGGGGAGATCGACTGTATCCCGGGGCTGGCGCTGTGACGGCTTT 3900
Db 3860 TTTGATGTGCAACAGGGGAGATCGACTGTATCCCGGGGCTGGCGCTGTGACGGCTTT 3919
QY 3901 CCGAGTGCATGACACAGAGCGACGAGAGGGCTGCCCGCTGTGTCTCCCGCCAGTTTC 3960
Db 3920 CCGAGTGCATGACACAGAGCGACGAGAGGGCTGCCCGCTGTGTCTCCCGCCAGTTTC 3979
QY 3961 CCTCGCGCGGGGTCTAGTGTGTGACCTGTGCGCTGGCTGGAGCGGAGGAGCACTGT 4020
Db 3980 CCTCGCGCGGGGTCTAGTGTGTGACCTGTGCGCTGGCTGGAGCGGAGGAGCACTGT 4039
QY 4021 CAGGACCGCTCAGACGAGGCGGACTGTGACGCCATCTGCCCTGCCCAACAGTTCCGGTGT 4080
Db 4040 CAGGACCGCTCAGACGAGGCTGACTGTGACGCCATCTGCCCTGCCCAACAGTTCCGGTGT 4099
QY 4081 GCGAGCGGCTAGTGTCTCATCAAAACAGAGTGTGACTCTTCTCCCGGACTGTATCGAC 4140
Db 4100 GCGAGCGGCTAGTGTCTCATCAAAACAGAGTGTGACTCTTCTCCCGGACTGTATCGAC 4159
QY 4141 GGTCTCGAGAGCTCATGTGTGAATCAACAAGCGCCCTCAGACGACAGCCGCGCCAC 4200
Db 4160 GGTCTCGAGAGCTCATGTGTGAATCAACAAGCGCCCTCAGACGACAGCCGCGCCAC 4219
QY 4201 AGCAGTGCATCTCGGCGCCGTCAATTTGGCATCATCTCTCTCTTCTGTCATGGGTGGTGC 4260
Db 4220 AGCAGTGCATCTCGGCGCCGTCAATTTGGCATCATCTCTCTCTTCTGTCATGGGTGGTGC 4279
QY 4261 TATTTTGTGTCACAGCGGTGGTGTGACAGGCTATGCGGGGGCAACGGGCGCTTCCG 4320
Db 4280 TATTTTGTGTCACAGCGGTGGTGTGACAGGCTATGCGGGGGCAACGGGCGCTTCCG 4339
QY 4321 CACGAGTATGTACGCGGACCCCGCATCGTGCCTCTCAATTTCTAGCCCGGGGGTTC 4380
Db 4340 CACGAGTATGTACGCGGACCCCGCATCGTGCCTCTCAATTTCTAGCCCGGGGGTTC 4399
QY 4381 CAGCATGGCCCTTTCACAGGCAATCGCATGCGGAAAGTCCATGATGAGTCCGTGAGCTG 4440
Db 4400 CAGCATGGCCCTTTCACAGGCAATCGCATGCGGAAAGTCCATGATGAGTCCGTGAGCTG 4459
QY 4441 ATGGGGGGCGGGGGGGTCCCTGTATACACCGGACACGTCACAGGGGCTTCGTC 4500
Db 4460 ATGGGGGGCGGGGGGGTCCCTCTTACACCGGAAACACGTCACAGGGGCTTCGTC 4519
QY 4501 AGCAGCTGTCTCAGACGAAAGCCAGCTGTATACCGCGCATCTGAAACCGCGCGCTTC 4560
Db 4520 AGCAGCTGTCTCAGACGAAAGCCAGCTGTATACCGCGCATCTGAAACCGCGCGCTTC 4579
QY 4561 CCGGCGACGAGCCCTCTCTGTACAAATGAGCAATGTTCTACTCTTCAAAATTCCTCGGCC 4620
Db 4580 CCGGCGACGAGCCCTCTCTGTACAAATGAGCAATGTTCTACTCTTCAAAATTCCTCGGCC 4639
QY 4621 ACTCGGAGCCGTACAGCCCTTACATTCGAGGAATGGCGCCCGACGACGCGCTTCG 4680
Db 4640 ACTCGGAGCCGTACAGCCCTTACATTCGAGGAATGGCGCCCGACGACGCGCTTCG 4699
QY 4681 AGCACCGAGCTGTGTGACGAGCTACAGCCCGCTGGAAGGCGGAGGCTACTTAC 4740
Db 4700 AGCACCGAGCTGTGTGACGAGCTACAGCCCGCTGGAAGGCGGAGGCTACTTAC 4759
QY 4741 CTGATTTGAACTCGGACTCAGACCCCTTATCCACCCCGACCCACGCGCCACGAGCAGTAC 4800
Db 4760 CTGATTTGAACTCGGACTCAGACCCCTTATCCACCCCGACCCACGCGCCACGAGCAGTAC 4819

QY 1381 GSCCTCAACGGCACCTCCGCAAGATCCTGGTGTGCGAGGACCTCGAGAGCCCGGAGCC 1440
Db 1429 CCGCTCAACGGCACCTCCGCAAGATCCTGGTGTGCGAGGACCTCGAGAGCCCGGAGCC 1488
QY 1441 ATCGCACTGCACCCCGTGATGGCCCTCATGTACTGGACAGATCGGGGAGAGAACCTTAAA 1500
Db 1489 ATCGCACTGCACCCCGTGATGGCCCTCATGTACTGGACAGATCGGGGAGAGAACCTTAAA 1548
QY 1501 ATCGAGTGTGCCAATTTGGATGGGAGAGAGCGGCGTGTGCTGCTCAATGCTCCCTCGGG 1560
Db 1549 ATCGAGTGTGCCAATTTGGATGGGAGAGAGCGGCGTGTGCTGCTCAATGCTCCCTCGGG 1608
QY 1561 TGGCCCAACGGCCTGGCCCTGGACCTGCAGAGGGGAGGAGCTTACTTGGGGAGAGCCCAAG 1620
Db 1609 TGGCCCAACGGCCTGGCCCTGGACCTGCAGAGGGGAGGAGCTTACTTGGGGAGAGCCCAAG 1668
QY 1621 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAAAGAGCGGAGCCCTCTCGAGGAC 1680
Db 1669 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAAAGAGCGGAGCCCTCTCGAGGAC 1728
QY 1681 AAGCTCCGCGACATTTTCGGGTTACGCTGTCTGGGGGACTTCACTACTGGACTGACTGG 1740
Db 1729 AAGCTCCGCGACATTTTCGGGTTACGCTGTCTGGGGGACTTCACTACTGGACTGACTGG 1788
QY 1741 CAGGCGCGCAGCATCGAGCGGTGCACAAGGTCAAGGCGCAGCGGAGCGTCACTCATTTGAC 1800
Db 1789 CAGGCGCGCAGCATCGAGCGGTGCACAAGGTCAAGGCGCAGCGGAGCGTCACTCATTTGAC 1848
QY 1801 CAGCTGCCCGACCTGATGGGGTCAAAGCTGTGAATGTGGCCAGGTCGTGGAAACCAAC 1860
Db 1849 CAGCTGCCCGACCTGATGGGGTCAAAGCTGTGAATGTGGCCAGGTCGTGGAAACCAAC 1908
QY 1861 CGGTGTGGGACAGAAACGGGGGTGACGACCTGTGCTTCTTCAACCCGACGCAACC 1920
Db 1909 CGGTGTGGGACAGAAACGGGGGTGACGACCTGTGCTTCTTCAACCCGACGCAACC 1968
QY 1921 CGGTGTGGCTGCCCATCGGCTGGAGCTGTGAGTGACATGAAGACCTGCATCGTGCT 1980
Db 1969 CGGTGTGGCTGCCCATCGGCTGGAGCTGTGAGTGACATGAAGACCTGCATCGTGCT 2028
QY 1981 GAGGCTTTCTTGGTCTTCAACAGCAGAGCCGCGCATCCACAGGATCTCCCTCGAGACCAAT 2040
Db 2029 GAGGCTTTCTTGGTCTTCAACAGCAGAGCCGCGCATCCACAGGATCTCCCTCGAGACCAAT 2088
QY 2041 AACACGAGCTGGCCATCCCGCTACGGGCGTCAAGGAGGCTCAGCCCTGGAATTGAT 2100
Db 2089 AACACGAGCTGGCCATCCCGCTACGGGCGTCAAGGAGGCTCAGCCCTGGAATTGAT 2148
QY 2101 GTGTCCAAACCAACATCTACTGGACAGACGTTCAGCCTGAAGACCATCAGCCGCGCCTTC 2160
Db 2149 GTGTCCAAACCAACATCTACTGGACAGACGTTCAGCCTGAAGACCATCAGCCGCGCCTTC 2208
QY 2161 ATGAACGGGAGCTCGGTGGAGCAGTGTGGAGTTTGGCCTTGAATACCCCGAGGGCATG 2220
Db 2209 ATGAACGGGAGCTCGGTGGAGCAGTGTGGAGTTTGGCCTTGAATACCCCGAGGGCATG 2268
QY 2221 GCCGTTGACTGGATGGGCAAGAACTCTACTGGGCGGACACTTGGGACCAACAGAAATCGAA 2280
Db 2269 GCCGTTGACTGGATGGGCAAGAACTCTACTGGGCGGACACTTGGGACCAACAGAAATCGAA 2328
QY 2281 GTGGCGGGCTGGACGGGCAAGTTCGGCAAGTCTCTGTGTGGAGGACTTGGACAAACCCG 2340
Db 2329 GTGGCGGGCTGGACGGGCAAGTTCGGCAAGTCTCTGTGTGGAGGACTTGGACAAACCCG 2388
QY 2341 AGGTGCTGGCCCTGGATCCCAACGAGGCTTACATCTACTGGACCGAGTGGGGCGGCAAG 2400
Db 2389 AGGTGCTGGCCCTGGATCCCAACGAGGCTTACATCTACTGGACCGAGTGGGGCGGCAAG 2448
QY 2401 CCGAGGATCGGGGCTTCTATGGACGGGACCAACTGATGACGCTGGTGGACAGAGTG 2460
Db 2449 CCGAGGATCGGGGCTTCTATGGACGGGACCAACTGATGACGCTGGTGGACAGAGTG 2508

QY 2461 GGCGGGGCAACGACCTCACCAITGACTACGCTGACAGCGCCTCTACTGAGACCGACCTG 2520
Db 2509 GGCGGGGCAACGACCTCACCAITGACTACGCTGACAGCGCCTCTACTGAGACCGACCTG 2568
QY 2521 GACACCAACATGATCGAGTCTGTCACCAACATGCTGGGTGAGGAGCGGTCGTGATTCGAG 2580
Db 2569 GACACCAACATGATCGAGTCTGTCACCAACATGCTGGGTGAGGAGCGGTCGTGATTCGAG 2628
QY 2581 GATCTCCGCGACCGGTTCTGGTCTGACGAGTACAGGATTTATCTACTGAGACAGACTGG 2640
Db 2629 GATCTCCGCGACCGGTTCTGGTCTGACGAGTACAGGATTTATCTACTGAGACAGACTGG 2688
QY 2641 AATCTGCACAGCATTTGAGCGGGCGACAAAGACTAGCGCGGAAACCGCACCTCATCCAG 2700
Db 2689 AATCTGCACAGCATTTGAGCGGGCGACAAAGACTAGCGCGGAAACCGCACCTCATCCAG 2748
QY 2701 GGCCACCTTGGACTTCTGTGATGGACATCCTGTGTGTTTCACTCTCCCGCAGGATGGCCTC 2760
Db 2749 GGCCACCTTGGACTTCTGTGATGGACATCCTGTGTGTTTCACTCTCCCGCAGGATGGCCTC 2808
QY 2761 AATGACTGTATGACAAACAGGCGAGTGTGGGAGCTGTGCTTGCCTATCCCGGCGGC 2820
Db 2809 AATGACTGTATGACAAACAGGCGAGTGTGGGAGCTGTGCTTGCCTATCCCGGCGGC 2868
QY 2821 CACCGCTGCGGCTCGGCTCACACTACACCTTGACCCAGCAGCGCAACTGACAGCCG 2880
Db 2869 CACCGCTGCGGCTCGGCTCACACTACACCTTGACCCAGCAGCGCAACTGACAGCCG 2928
QY 2881 CCCACCACTTTCTGTCTTTCAGCCAGAAATCTGCCATCAGTCGGATGATCCCGAGCAC 2940
Db 2929 CCCACCACTTTCTGTCTTTCAGCCAGAAATCTGCCATCAGTCGGATGATCCCGAGCAC 2988
QY 2941 CAGCACAGCCCGGATCTCATCTCTGCCCTGATGGATGGAGAACTCAAGACCATCCAC 3000
Db 2989 CAGCACAGCCCGGATCTCATCTCTGCCCTGATGGATGGAGAACTCAAGACCATCCAC 3048
QY 3001 TATGACCCACCTGGACAAGTTTCTATCTACTGGTGGATGGGCGCCAGAACTCAAGCGAGCC 3060
Db 3049 TATGACCCACCTGGACAAGTTTCTATCTACTGGTGGATGGGCGCCAGAACTCAAGCGAGCC 3108
QY 3061 AAGGACGAGCGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAAAGGCCAAAACCCAGAC 3120
Db 3109 AAGGACGAGCGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAAAGGCCAAAACCCAGAC 3168
QY 3121 AGGACGACCCACAGACTCAGCATCGACATCTACAGCGGACACTGTTCTGACGCTGCGAG 3180
Db 3169 AGGACGACCCACAGACTCAGCATCGACATCTACAGCGGACACTGTTCTGACGCTGCGAG 3228
QY 3181 GCCACCAATACATCAACGCTCCAGGCTGAGCGGGGAAAGCCATGGGGGTGGTCTGCGT 3240
Db 3229 GCCACCAATACATCAACGCTCCAGGCTGAGCGGGGAAAGCCATGGGGGTGGTCTGCGT 3288
QY 3241 GGGGACCGGACAAAGCCAGGGCCATCGTGTCAACCGGAGCGAGGGTACTCTGTACTTC 3300
Db 3289 GGGGACCGGACAAAGCCAGGGCCATCGTGTCAACCGGAGCGAGGGTACTCTGTACTTC 3348
QY 3301 ACCAACATGAGGAGCGGGGAGCCAAAGATCGAAACCGCGAGCCCTGGAGCGGACCGAGCGC 3360
Db 3349 ACCAACATGAGGAGCGGGGAGCCAAAGATCGAAACCGCGAGCCCTGGAGCGGACCGAGCGC 3408
QY 3361 GAGGTCTCTTACACACCGGCTCATCCGCTGTGCGCTGTGGCTGTGGTGGACAAACACTG 3420
Db 3409 GAGGTCTCTTACACACCGGCTCATCCGCTGTGCGCTGTGGCTGTGGTGGACAAACACTG 3468
QY 3421 GGCAGGCTGTCTGGGTGGAGCGGACCTGAAAGCGCATTTGAGAGCTGTGACTGTGAGG 3480
Db 3469 GGCAGGCTGTCTGGGTGGAGCGGACCTGAAAGCGCATTTGAGAGCTGTGACTGTGAGG 3528
QY 3481 GCCAACCGCTGACCTTGGAGGAGCGCAACATCTGTGAGCCTCTGGGCTGACCATCTCTT 3540
Db 3529 GCCAACCGCTGACCTTGGAGGAGCGCAACATCTGTGAGCCTCTGGGCTGACCATCTCTT 3588
QY 3541 GGCAAGCATCTCTACTGGATCGACCCGCGCAGCAGCATGATCGAGCGTGTGGAGAGACC 3600

[illegible]

Db	3409	GAGTCTCTTTCACCA	CGGGCCTCATCCGGCCTTGTTGGCCCTGTGTGGTGGAGACAACACACTG	3468
Qy	3421	GGCAAGCTGTCTCGGTGGAC	CGCGGACCTGAAGCGCAATTGAGAGCTGTGACCTGTTCAGGG	3480
Db	3469	GGCAAGCTGTCTCGGTGGAC	CGCGGACCTGAGGGCAATTGAGAGCTGTGACCTGTTCAGGG	3528
Qy	3481	GCCAACCGCTTGACCCCTGGAGGAC	CGCAACATGTGTGAGCTCTTGGGCCCTTGACCATCTCTT	3540
Db	3529	GCCAACCGCTTGACCCCTGGAGGAC	CGCAACATGTGTGAGCTCTTGGGCCCTTGACCATCTCTT	3588
Qy	3541	GGCAAGCATCTCTACTGGATCGAC	CGCGGACCGGATGATCGAGCTGTGGAGAGAC	3600
Db	3589	GGCAAGCATCTCTACTGGATCGAC	CGCGGACCGGATGATCGAGCTGTGGAGAGAC	3648
Qy	3601	ACCGGGGACAAAGCGGACCTCGCAT	CCAGGGCGGTGTCGCCCACTTCATCTGGCATCCATGCA	3660
Db	3649	ACCGGGGACAAAGCGGACCTCGCAT	CCAGGGCGGTGTCGCCCACTTCATCTGGCATCCATGCA	3708
Qy	3661	GTGAGGAGATGACCTTGAGGAGATT	CTCAGCCACCATGTGCGCTGTGACATTAATGTGTGGC	3720
Db	3709	GTGAGGAGATGACCTTGAGGAGATT	CTCAGCCACCATGTGCGCTGTGACATTAATGTGTGGC	3768
Qy	3721	TGCTCCACCATCTGTATTGCGAAGG	TGATGGGACACACAGGTGCTCATGCCAGTCCAC	3780
Db	3769	TGCTCCACCATCTGTATTGCGAAGG	TGATGGGACACACAGGTGCTCATGCCAGTCCAC	3828
Qy	3781	CTCGTGCTCTGTCAGAACCTGTGAC	CTGTGGAGAGCGGCCACTCTCTCCCGGACACG	3840
Db	3829	CTCGTGCTCTGTCAGAACCTGTGAC	CTGTGGAGAGCGGCCACTCTCTCCCGGACACG	3888
Qy	3841	TTTGATGTGTCACAGGGGAGATCGAC	TGTATCCCGGGGCTGTGGCGCTGTGACGGCTTT	3900
Db	3889	TTTGATGTGTCACAGGGGAGATCGAC	TGTATCCCGGGGCTGTGGCGCTGTGACGGCTTT	3948
Qy	3901	CCGAGTGCATGATCAGAGCGACGAG	AGGGCTGCCCGTGTGCTCCGCGCGCCAGTTC	3960
Db	3949	CCGAGTGCATGATCAGAGCGACGAG	AGGGGTGCCCCGTGTGCTCCGCGCGCCAGTTC	4008
Qy	3961	CCCTGCCGCGGGGTCAGTGTGTGGAC	CTCGCTGCGTCGACGCGAGGCGAGGACACTGT	4020
Db	4009	CCCTGCCGCGGGGTCAGTGTGTGGAC	CTCGCTGCGTCGACGCGAGGCGAGGACACTGT	4068
Qy	4021	CAGGACCGCTCAGACGAGGGCGACT	GTGACGCCATCTGCTTCCGTCGCCAACAGTTTCCGGTGT	4080
Db	4069	CAGGACCGCTCAGACGAGGGCGACT	GTGACGCCATCTGCTTCCGTCGCCAACAGTTTCCGGTGT	4128
Qy	4081	CGGAGCGGCGAGTGTCTCATCAAA	CAGAGTGTGACTCTTCCCGGACTGTATCGAC	4140
Db	4129	CGGAGCGGCGAGTGTCTCATCAAA	CAGAGTGTGACTCTTCCCGGACTGTATCGAC	4188
Qy	4141	GGCTCCAGCAGCTCATGTGTGAAAT	CACACGCGGCTCAGACGACACCGCGGCCAC	4200
Db	4189	GGCTCCAGCAGCTCATGTGTGAAAT	CACACGCGGCTCAGACGACACCGCGGCCAC	4248
Qy	4201	AGCAGTGCCATTCGGGCCCGCTCAT	TGGCATCATCTCTCTCTTCTGTCATGGGTGGTGC	4260
Db	4249	AGCAGTGCCATTCGGGCCCGCTCAT	TGGCATCATCTCTCTCTTCTGTCATGGGTGGTGC	4308
Qy	4261	TATTTTGTGTGCAGCGCGTGTGTGC	ACGCTATGCGGGGCGCAACCGGGCCCTTCCCG	4320
Db	4309	TATTTTGTGTGCAGCGCGTGTGTGC	ACGCTATGCGGGGCGCAACCGGGCCCTTCCCG	4368
Qy	4321	CACAGTATGTCAGCGGACCCCGCAG	TCGTCGCTCAATTTTCATGCCCCCGGGGTTC	4380
Db	4369	CACAGTATGTCAGCGGACCCCGCAG	TCGTCGCTCAATTTTCATGCCCCCGGGGTTC	4428
Qy	4381	CAGCATGGCCCCCTTTCAGAGCATCG	CATGCGGAAAGTCCATGATGAGCTCCGTGAGCCTG	4440
Db	4429	CAGCATGGCCCCCTTTCAGAGCATCG	CATGCGGAAAGTCCATGATGAGCTCCGTGAGCCTG	4488
Qy	4441	ATGGGGGCGCGGGGTGCCCTGTACGA	CCGGAAACGTCACAGGGGCTCTGTC	4500
Db	4489	ATGGGGGCGCGGGGTGCCCTGTACGA	CCGGAAACGTCACAGGGGCTCTGTC	4548

Db 1100 CTGCTGTCGGCCGGCGGACGGAAGCTACGAGGATCTCGTGGACACGCGGACTTCACC 1159
QY 1141 GACATCGTGTGTCAGGTGGACGACATCCGGCAGCCATGGCCATCGACTAGACCCGCTA 1200
Db 1160 GACATCGTGTGTCAGGTGGACGACATCCGGCAGCCATGGCCATCGACTAGACCCGCTA 1219
QY 1201 GAGGGCTATGTCTACTGGACAGATGACGAGGTGGGGCCATCCGACAGGGCGTACTCGAC 1260
Db 1220 GAGGGCTATGTCTACTGGACAGATGACGAGGTGGGGCCATCCGACAGGGCGTACTCGAC 1279
QY 1261 GGGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGTTC 1320
Db 1280 GGGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGTTC 1339
QY 1321 GACTGGGTGGCCGAAACCTTACTTGGACCGACACGCGGACCGGATCGAGGTGACG 1380
Db 1340 GACTGGGTGGCCGAAACCTTACTTGGACCGACACGCGGACCGGATCGAGGTGACG 1399
QY 1381 CGCCTCAACGGCACCTCCCGCAAGATCCTGGTGTGGAGGACCTGGACGAGCCCGGAGCC 1440
Db 1400 CGCCTCAACGGCACCTCCCGCAAGATCCTGGTGTGGAGGACCTGGACGAGCCCGGAGCC 1459
QY 1441 ATCGCATGCAACCCGATGAGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAAA 1500
Db 1460 ATCGCATGCAACCCGATGAGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAAA 1519
QY 1501 ATCGAGTGTGCAACTTGGATGGGACGAGGCGGTGTGTGGTCAATGCTTCCCTCGG 1560
Db 1520 ATCGAGTGTGCAACTTGGATGGGACGAGGCGGTGTGTGGTCAATGCTTCCCTCGG 1579
QY 1561 TGGCCCAACGGCCTTGGCCTGGACCTGAGAGGGAGAGCTTACTGGGAGAGCCCAAG 1620
Db 1580 TGGCCCAACGGCCTTGGCCTGGACCTGAGAGGGAGAGCTTACTGGGAGAGCCCAAG 1639
QY 1621 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAGAGGGAGAGCCCTCTGGAGAC 1680
Db 1640 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAGAGGGAGAGCCCTCTGGAGAC 1699
QY 1681 AAGCTCCGCGACATTTTCGGGTTTCAAGTGTGGGGAGCTTACTTACTGGAGTACTGG 1740
Db 1700 AAGCTCCGCGACATTTTCGGGTTTCAAGTGTGGGGAGCTTACTTACTGGAGTACTGG 1759
QY 1741 CAGCGCCGACATCGAGCGGTGCACAAGTCAAGGCGCAGCGGAGCTCATCATTTGAC 1800
Db 1760 CAGCGCCGACATCGAGCGGTGCACAAGTCAAGGCGCAGCGGAGCTCATCATTTGAC 1819
QY 1801 CAGCTGCCCGACCTGATGGGCTCAAAGCTGTGAATGTGGCCAAAGTGTGGAAACCAAC 1860
Db 1820 CAGCTGCCCGACCTGATGGGCTCAAAGCTGTGAATGTGGCCAAAGTGTGGAAACCAAC 1879
QY 1861 CCGTGTGGGACAGAAACGGGGGTGCAGCCACTGTGTCTTCTTCAACCCCGACCAAC 1920
Db 1880 CCGTGTGGGACAGAAACGGGGGTGCAGCCACTGTGTCTTCTTCAACCCCGACCAAC 1939
QY 1921 CGGTGTGCTGCCCATCGGCTCGAGCTGTGAGTGCATGAAGACCTGTGATCGTGCCT 1980
Db 1940 CGGTGTGCTGCCCATCGGCTCGAGCTGTGAGTGCATGAAGACCTGTGATCGTGCCT 1999
QY 1981 GAGGCTTTCTTGGTCTTCAACAGCAGAGCCCGCCATCCACAGGATCTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTTTCTTGGTCTTCAACAGCAGAGCCCGCCATCCACAGGATCTCCCTCGAGACCAAT 2059
QY 2041 AACAAAGAGTGGGCTATCCCGCTCAACGGGCTCAAGAGGCTCAGGCTGGATTTGAT 2100
Db 2060 AACAAAGAGTGGGCTATCCCGCTCAACGGGCTCAAGAGGCTCAGGCTGGATTTGAT 2119
QY 2101 GTGTCCCAACACATCTACTGACAGAGCTGAGCTGAGAGCCATCAGCGCGCTTC 2160
Db 2120 GTGTCCCAACACATCTACTGACAGAGCTGAGCTGAGAGCCATCAGCGCGCTTC 2179
QY 2161 ATGAACGGGAGCTCGGTGGAGACGCTGTGTGAGTTTGGGCTTGTACTACCCCGAGGGATG 2220

Db 2180 ATGAACGGGAGCTCGGTGGAGACGCTGTGTGAGTTTGGCCCTTGACTACCCCGAGGCGATG 2239
QY 2221 GCCGTTCACATGGATGGGCAAGAACCTTACTTGGGCCGACACTGGGACCAACAGAAATCGAA 2280
Db 2240 GCCGTTCACATGGATGGGCAAGAACCTTACTTGGGCCGACACTGGGACCAACAGAAATCGAA 2299
QY 2281 GTGCGCGGCTGGACGCGGCGATTCGCGCAAGTCCCTCGTGTGGAGGAGCTTGGACAAACCCG 2340
Db 2300 GTGCGCGGCTGGACGCGGCGATTCGCGCAAGTCCCTCGTGTGGAGGAGCTTGGACAAACCCG 2359
QY 2341 AGTTCGCTGGCCTGGATTCGCAACAGGGCTACATCTACTTGGACCGAGTGGGGCGGCAAG 2400
Db 2360 AGTTCGCTGGCCTGGATTCGCAACAGGGCTACATCTACTTGGACCGAGTGGGGCGGCAAG 2419
QY 2401 CCGAGGATCGTGGGGCTTTCATGGACGGGACCAACCTGCAATGATGACGCTGGTGGACAGGTG 2460
Db 2420 CCGAGGATCGTGGGGCTTTCATGGACGGGACCAACCTGCAATGATGACGCTGGTGGACAGGTG 2479
QY 2461 GGC CGGCGCCAAACGACCTTCAACATTTGACTAGCTGACGAGCGCTTCTACTTGGACCGACCTG 2520
Db 2480 GGC CGGCGCCAAACGACCTTCAACATTTGACTAGCTGACGAGCGCTTCTACTTGGACCGACCTG 2539
QY 2521 GACACCAACATGATGAGTGTGCTGCAACATGCTGGGTCAAGAGCGGGTGTGATTTGCCGAC 2580
Db 2540 GACACCAACATGATGAGTGTGCTGCAACATGCTGGGTCAAGAGCGGGTGTGATTTGCCGAC 2599
QY 2581 GATCTCCGCGACCCGTTCCGTTCTGACGAGTACAGCGATTTATATCTACTTGGACAGACTGG 2640
Db 2600 GATCTCCGCGACCCGTTCCGTTCTGACGAGTACAGCGATTTATATCTACTTGGACAGACTGG 2659
QY 2641 AATCTGCAACGACCTTGAAGCGGCGCGCAACAGACTAGCGGCGCGGAAACCGCACCTTATCCAG 2700
Db 2660 AATCTGCAACGACCTTGAAGCGGCGCGCAACAGACTAGCGGCGCGGAAACCGCACCTTATCCAG 2719
QY 2701 GGCACCTGAGCTTCTGATGGACATCTCGTGTTCATCTCTCCGCGCAGGATGGGCTC 2760
Db 2720 GGCACCTGAGCTTCTGATGGACATCTCGTGTTCATCTCTCCGCGCAGGATGGGCTC 2779
QY 2761 AATGACTGTATGCAACAAACGGGCGAGTGGGCGAGCTGTGCTTGCCTCCCTCCCGGGCGC 2820
Db 2780 AATGACTGTATGCAACAAACGGGCGAGTGGGCGAGCTGTGCTTGCCTCCCTCCCGGGCGC 2839
QY 2821 CACCGCTGCGGCTGCGCTCACACTACACCTTGACCCCGACGAGCCGCAACTGACGCGCG 2880
Db 2840 CACCGCTGCGGCTGCGCTCACACTACACCTTGACCCCGACGAGCCGCAACTGACGCGCG 2899
QY 2881 CCAACCACTTCTGCTGTTCAGCCAGAAATCTGCCATCTAGTTCGATGATCCCGGACGAC 2940
Db 2900 CCAACCACTTCTGCTGTTCAGCCAGAAATCTGCCATCTAGTTCGATGATCCCGGACGAC 2959
QY 2941 CAGCAGCGCGGATCTCATCTGCTGCCCTGCATGGACTGAGGACGTCGAAGCCATCGAC 3000
Db 2960 CAGCAGCGCGGATCTCATCTGCTGCCCTGCATGGACTGAGGACGTCGAAGCCATCGAC 3019
QY 3001 TATGACCCCACTGGACAAAGTTTCATCTACTGGTGGATGGGCGCCAGAAACATCAAGCGAGCC 3060
Db 3020 TATGACCCCACTGGACAAAGTTTCATCTACTGGTGGATGGGCGCCAGAAACATCAAGCGAGCC 3079
QY 3061 AAGSACGACGGGACCCAGCCCTTTGTTTGAACCTCTCTGAGCCAAAGGCCAAAACCCAGAC 3120
Db 3080 AAGSACGACGGGACCCAGCCCTTTGTTTGAACCTCTCTGAGCCAAAGGCCAAAACCCAGAC 3139
QY 3121 AGGAGCGCCACGACCTCAGCATCGACATCTAGCGCGGACACTGTCTTGGACCTGGAG 3180
Db 3140 AGGAGCGCCACGACCTCAGCATCGACATCTAGCGCGGACACTGTCTTGGACCTGGAG 3199
QY 3181 GCCACCAATACCACTCAACGCTCAGCGGGGAGCCATGGGGGTGGTCTGGT 3240
Db 3200 GCCACCAATACCACTCAACGCTCAGCGGGGAGCCATGGGGGTGGTCTGGT 3259
QY 3241 GGGGACCGCGACAAAGCCCGAGGGCCATCGTGTCAACCGCGGAGCGAGGGTACTCTGTACTTC 3300
Db 3260 GGGGACCGCGACAAAGCCCGAGGGCCATCGTGTCAACCGCGGAGCGAGGGTACTCTGTACTTC 3319

3001 TATGACCCACTGGACAAGTTTCTACTTACTGGGTGGATGGCGCCAGAAATCAATCAAGCGAGCC 3060
3020 TATGACCCACTGGACAAGTTTCTACTTACTGGGTGGATGGCGCCAGAAATCAATCAAGCGAGCC 3079
3061 AAGGACGACGGAGCCAGCCCTTTGTTGACCTCTCTGAGCCAAAGGCGAAACCCAGAC 3120
3080 AAGGACGACGGAGCCAGCCCTTTGTTGACCTCTCTGAGCCAAAGGCGAAACCCAGAC 3139
3121 AGGACGACCCAGGACCTCAGATCGACATCTACAGCCGGAACATGTTTCTGACGTCGAG 3180
3140 AGGACGACCCAGGACCTCAGATCGACATCTACAGCCGGAACATGTTTCTGACGTCGAG 3199
3181 GCCACCAATACCATCAAGCTTCAAGGCTGAGCGGGGAAGCCATGGGGGTGTTGCTGGCT 3240
3200 GCCACCAATACCATCAAGCTTCAAGGCTGAGCGGGGAAGCCATGGGGGTGTTGCTGGCT 3259
3241 GGGGACCGCGACAAGCCCAAGGCGCATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 3300
3260 GGGGACCGCGACAAGCCCAAGGCGCATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 3319
3301 ACCAATGACGAGCCGCGGACCAAGATCGAAGCGCGAGCCCTGGAACGCGAGCGCC 3360
3320 ACCAATGACGAGCCGCGGACCAAGATCGAAGCGCGAGCCCTGGAACGCGAGCGCC 3379
3361 GAGGTCTCTTCAACACCGGCTCATCCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTG 3420
3380 GAGGTCTCTTCAACACCGGCTCATCCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTG 3439
3421 GCGAAGCTGTTCTGGGTGGACCGCGACTGAGGCGCATTTGAGAGCTGTGACCTGTCAGGG 3480
3440 GCGAAGCTGTTCTGGGTGGACCGCGACTGAGGCGCATTTGAGAGCTGTGACCTGTCAGGG 3499
3481 GCGAAGCTGTTCTGGGTGGACCGCGACTGAGGCGCATTTGAGAGCTGTGACCTGTCAGGG 3540
3500 GCGAAGCTGTTCTGGGTGGACCGCGACTGAGGCGCATTTGAGAGCTGTGACCTGTCAGGG 3559
3541 GCGAAGCTGTTCTGGGTGGACCGCGACTGAGGCGCATTTGAGAGCTGTGACCTGTCAGGG 3600
3560 GCGAAGCTGTTCTGGGTGGACCGCGACTGAGGCGCATTTGAGAGCTGTGACCTGTCAGGG 3619
3601 ACCGGGACAAGCGGACTCGCATCCAGGCGCTGTCGCGCCACCTCACTGGCATCCATGCA 3660
3620 ACCGGGACAAGCGGACTCGCATCCAGGCGCTGTCGCGCCACCTCACTGGCATCCATGCA 3679
3661 GTGAGGAGTCAAGCTGAGGAGTCTCAGCCACCATGTGCGCGTGGACAAATGGTGGC 3720
3680 GTGAGGAGTCAAGCTGAGGAGTCTCAGCCACCATGTGCGCGTGGACAAATGGTGGC 3739
3721 TGCTCCCACTCTGATTGCGAAGGTCATGGGACCAACAGCGTGTGATGCGGAGTCCAC 3780
3740 TGCTCCCACTCTGATTGCGAAGGTCATGGGACCAACAGCGTGTGATGCGGAGTCCAC 3799
3781 CTGCTGCTCTGCAAACTGCTGACCTGTGAGAGCGCGCCACCTGCTCCCGGACGAG 3840
3800 CTGCTGCTCTGCAAACTGCTGACCTGTGAGAGCGCGCCACCTGCTCCCGGACGAG 3859
3841 TTTGATGTCGCAAGGAGATCGACTGATATCCCGGGGCTGGCGCTGTGACGGCTTT 3900
3860 TTTGATGTCGCAAGGAGATCGACTGATATCCCGGGGCTGGCGCTGTGACGGCTTT 3919
3901 CCCGAGTGCATGATGACGAGGAGCGAGGAGGCTGCGCGCTGTGCTCCCGCGCCAGTTC 3960
3920 CCCGAGTGCATGATGACGAGGAGCGAGGAGGCTGCGCGCTGTGCTCCCGCGCCAGTTC 3979
3961 CCCTGCGCGGGGTCAAGTGTGAGACTGCGCTGCGCTGCGAGCGGAGGAGGACTGT 4020
3980 CCCTGCGCGGGGTCAAGTGTGAGACTGCGCTGCGCTGCGAGCGGAGGAGGACTGT 4039
4021 CAGGACCGCTCAGACGAGGCGGACTGTGACCGGCTGCGCTGCGGACCGGAGGAGGACTGT 4080
4040 CAGGACCGCTCAGACGAGGCGGACTGTGACCGGCTGCGCTGCGGACCGGAGGAGGACTGT 4099

4081 GCGAGCGCCAGTGTGCTCTCATCAAAACAGCAGTGCAGCTCTCTCCCGACTGTATCGAC 4140
4100 GCGAGCGCCAGTGTGCTCTCATCAAAACAGCAGTGCAGCTCTCTCCCGACTGTATCGAC 4159
4141 GGTCCCGACGAGCTCATGTGTGAATCAACAAAGCGCCCTCAGACGACGAGCCGCCAC 4200
4160 GGTCCCGACGAGCTCATGTGTGAATCAACAAAGCGCCCTCAGACGACGAGCCGCCAC 4219
4201 AGCAGTGCATCGGCGCGCTCATTTGGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4260
4220 AGCAGTGCATCGGCGCGCTCATTTGGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4279
4261 TATTTGTGTGCCAGCGCTGTGTCAGCGCTATGCGGGGCGCAACGAGGCGCTTCGCG 4320
4280 TATTTGTGTGCCAGCGCTGTGTCAGCGCTATGCGGGGCGCAACGAGGCGCTTCGCG 4339
4321 CACGAGTATGTACGCGGAGCCCCGCACTGCGCCCTCAATTTCAATGAGCCGCGGCGGTTC 4380
4340 CACGAGTATGTACGCGGAGCCCCGCACTGCGCCCTCAATTTCAATGAGCCGCGGCGGTTC 4399
4381 CAGCATGCGCCCTTTCACAGGCGATCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCCTG 4440
4400 CAGCATGCGCCCTTTCACAGGCGATCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCCTG 4459
4441 ATGGGGGCGCGGGCGGGTGCCTCTGTACGACCGGAAACACGTCACAGGGGCTTCGTC 4500
4460 ATGGGGGCGCGGGCGGGTGCCTCTGTACGACCGGAAACACGTCACAGGGGCTTCGTC 4519
4501 AGCAGCTGCTCCAGCAAGAGGCGCACTGCTGTACCGCGCGATCTCTGAAACCGCGGCGCTTC 4560
4520 AGCAGCTGCTCCAGCAAGAGGCGCACTGCTGTACCGCGCGATCTCTGAAACCGCGGCGCTTC 4579
4561 CCGGCGACGAGCCCTCCCTGTCAACATGAGCACTGTTCTACTCTTCAAAATTCGCGGC 4620
4580 CCGGCGACGAGCCCTCCCTGTCAACATGAGCACTGTTCTACTCTTCAAAATTCGCGGC 4639
4621 ACTCGAGACCGTACAGGCGCTACATCATTCAGGAATGGCGCGCCCGACGAGCGCTTCG 4680
4640 ACTCGAGACCGTACAGGCGCTACATCATTCAGGAATGGCGCGCCCGACGAGCGCTTCG 4699
4681 AGCAACGAGTGTGTGACGAGCTACAGCGCCAGCGCTGGAAGGCGCAAGTACTAC 4740
4700 AGCAACGAGTGTGTGACGAGCTACAGCGCCAGCGCTGGAAGGCGCAAGTACTAC 4759
4741 CTGATTTGAACTCGGACTCAGACCGCTTATCCACCGCCCGACCGCCCGACGAGCGTAC 4800
4760 CTGATTTGAACTCGGACTCAGACCGCTTATCCACCGCCCGACCGCCCGACGAGCGTAC 4819
4801 CTGCGCGGAGGACAGCTGCGCGCGCTCGCGCGCCAGCGAGGAGCTACTTCCATCTC 4860
4820 CTGCGCGGAGGACAGCTGCGCGCGCTCGCGCGCCAGCGAGGAGCTACTTCCATCTC 4879
4861 TTCCCGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 4920
4880 TTCCCGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 4939
4921 TTCTCTGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 4980
4940 TTCTCTGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 4999
4981 TTAATA 5040
5000 TTAATA 5059
5041 GGCAGGCGTGGAGAACTTTGTA 5063
5060 GGCAGGCGTGGAGAACTTTGTA 5082

RESULT 12

US-10-477-238A-2

; Sequence 2, Application US/10477238A

; Publication No. US20040221326A1

Db 1760 CAGCGCCGAGCATCGAGCGGTGCAAGAGTCAAGGCGAGCGGAGCGTCACTATTGAC 1819
Qy 1801 CAGCTGCCCGACCTGATCGGGGTCAAAGCTGTGAATGTGGCCAAAGGTGCTGGAAACCAAC 1860
Db 1820 CAGCTGCCCGACCTGATCGGGGTCAAAGCTGTGAATGTGGCCAAAGGTGCTGGAAACCAAC 1879
Qy 1861 CCGTGTGGCGGACAGGAACGGGGGTGCAGCCACCTGTGCTTCTTCAACCCCGACCAAC 1920
Db 1880 CCGTGTGGCGGACAGGAACGGGGGTGCAGCCACCTGTGCTTCTTCAACCCCGACCAAC 1939
Qy 1921 CCGTGTGGCTGCCCATCGGCTGTGAGCTGTGAGTGAATGAAGACCTGTGATGCTGCT 1980
Db 1940 CCGTGTGGCTGCCCATCGGCTGTGAGCTGTGAGTGAATGAAGACCTGTGATGCTGCTGCT 1999
Qy 1981 GAGGCTTCTTGTGCTTCAACAGCAGAGCGCCCATCCACAGATCTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTTCTTGTGCTTCAACAGCAGAGCGCCCATCCACAGATCTCCCTCGAGACCAAT 2059
Qy 2041 AACACGAGCTGGCCATCCCGCTCAACGGGCGGTCAAGGAGGCTCAGCCCTCGACTTTGAT 2100
Db 2060 AACACGAGCTGGCCATCCCGCTCAACGGGCGGTCAAGGAGGCTCAGCCCTCGACTTTGAT 2119
Qy 2101 GTGTCCAAACAACCATCTAATGGAAGACGCTGAGCTGGAAGACCATCAGCGCGCTTC 2160
Db 2120 GTGTCCAAACAACCATCTAATGGAAGACGCTGAGCTGGAAGACCATCAGCGCGCTTC 2179
Qy 2161 ATGACGGGAGCTGGTGGAGCAGCTGTGGAGTTGGCTTGTACTACCCGAGGGCATG 2220
Db 2180 ATGACGGGAGCTGGTGGAGCAGCTGTGGAGTTGGCTTGTACTACCCGAGGGCATG 2239
Qy 2221 GCCGTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAACAGAACTGAA 2280
Db 2240 GCCGTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAACAGAACTGAA 2299
Qy 2281 GTGGCGGGCTGGACGGGAGCTTCCGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCG 2340
Db 2300 GTGGCGGGCTGGACGGGAGCTTCCGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCG 2359
Qy 2341 AGGTGCTGGCCCTGGATCCCAACAGGGCTACATCTACTGGACCGAGTGGGGCGGAAG 2400
Db 2360 AGGTGCTGGCCCTGGATCCCAACAGGGCTACATCTACTGGACCGAGTGGGGCGGAAG 2419
Qy 2401 CCGAGGATCTGGCGGCTTCAATGACGGGACCAACTGCATGACGCTGTGGAACAAGGTG 2460
Db 2420 CCGAGGATCTGGCGGCTTCAATGACGGGACCAACTGCATGACGCTGTGGAACAAGGTG 2479
Qy 2461 GGCGGGCCAAACGACCTCAACATTGACTACGCTGACAGCGCTTCTACTGACCGACCTG 2520
Db 2480 GGCGGGCCAAACGACCTCAACATTGACTACGCTGACAGCGCTTCTACTGACCGACCTG 2539
Qy 2521 GACACCAATGATCGAGTCTGTCNAACATGCTGGGTGAGGAGCGGCTGATTTGCCGAC 2580
Db 2540 GACACCAATGATCGAGTCTGTCNAACATGCTGGGTGAGGAGCGGCTGATTTGCCGAC 2599
Qy 2581 GATCTCCCGCACCGCTTCCGCTGTGACGAGTACAGCGATTAATATCTACTGACAGACTGG 2640
Db 2600 GATCTCCCGCACCGCTTCCGCTGTGACGAGTACAGCGATTAATATCTACTGACAGACTGG 2659
Qy 2641 AATCTGCAACAGATTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCTCATCCAG 2700
Db 2660 AATCTGCAACAGATTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCTCATCCAG 2719
Qy 2701 GGCCACCTGGACTTGGTATGAGATCTCTGTGTGTTCACTCTCCCGCAGAGTGGCTC 2760
Db 2720 GGCCACCTGGACTTGGTATGAGATCTCTGTGTGTTCACTCTCCCGCAGAGTGGCTC 2779
Qy 2761 AATGACTGTATGCAACAACCGGCGAGTGTGGCAGCTGTGCTTGGCATCCCGGGGCG 2820
Db 2780 AATGACTGTATGCAACAACCGGCGAGTGTGGCAGCTGTGCTTGGCATCCCGGGGCG 2839
Qy 2821 CACCGCTCGGCTGGCTTCACTTACACCTTGAACCCAGACCGCAACTGACGCGCG 2880

Db 2840 CACCGCTCGGCTGGCTTCACTTACACCTTGAACCCAGCAGCCGCAACTGACGCGG 2899
Qy 2881 CCCACCACTTCTTGTCTTTCAGCCAGAAATCTGCCATCAGTTCGGATCATCCGACGAC 2940
Db 2900 CCCACCACTTCTTGTCTTTCAGCCAGAAATCTGCCATCAGTTCGGATCATCCGACGAC 2959
Qy 2941 CAGCACAGCCCGGATCTCATCTCTGCCCCTGATGAGTCTGAGGAACTCAAGACCATCGAC 3000
Db 2960 CAGCACAGCCCGGATCTCATCTCTGCCCCTGATGAGTCTGAGGAACTCAAGACCATCGAC 3019
Qy 3001 TATGACCACTCTGAGCAAGATTCTATCTGGGTGAGTGGGCGCAGAACTCAAGGAGCC 3060
Db 3020 TATGACCACTCTGAGCAAGATTCTATCTGGGTGAGTGGGCGCAGAACTCAAGGAGCC 3079
Qy 3061 AAGGACGAGCGGAGCCAGCCCTTGTGTTTGAACCTCTCTGAGCCAGGCGCAAAACCCAGAC 3120
Db 3080 AAGGACGAGCGGAGCCAGCCCTTGTGTTTGAACCTCTCTGAGCCAGGCGCAAAACCCAGAC 3139
Qy 3121 AGGACGAGCCCAAGACCTCTGAGCATCGACATCTTACAGCCGACACTGTTCTTGGACGTGCGAG 3180
Db 3140 AGGACGAGCCCAAGACCTCTGAGCATCGACATCTTACAGCCGACACTGTTCTTGGACGTGCGAG 3199
Qy 3181 GCCACCAATACCATCAAGTCCACAGGCTGAGCGGGGAAAGCCATGCGGGTGTGCTGGT 3240
Db 3200 GCCACCAATACCATCAAGTCCACAGGCTGAGCGGGGAAAGCCATGCGGGTGTGCTGGT 3259
Qy 3241 GGGGACCGGACGAGCCAGGCGCATCTGCTGCTCAACGGGAGCGGAGGCTGCTGACTTC 3300
Db 3260 GGGGACCGGACGAGCCAGGCGCATCTGCTGCTCAACGGGAGCGGAGGCTGCTGACTTC 3319
Qy 3301 ACCAACATGACGAGCCGCGGACGCAAGATCCGAGCGCCCTTGGACGCGACCCAGGCGC 3360
Db 3320 ACCAACATGACGAGCCGCGGACGCAAGATCCGAGCGCCCTTGGACGCGACCCAGGCGC 3379
Qy 3361 GAGTCTCTTTCACACACCGGCTCATTCGCTGTGGCCCTGTGGTGTGGACAAACACTG 3420
Db 3380 GAGTCTCTTTCACACACCGGCTCATTCGCTGTGGCCCTGTGGTGTGGACAAACACTG 3439
Qy 3421 GGCAGGCTGTTCTGGGTGAGCGGACCTGAAGCGCATTTGAGAGCTGTGACTGTGACGG 3480
Db 3440 GGCAGGCTGTTCTGGGTGAGCGGACCTGAAGCGCATTTGAGAGCTGTGACTGTGACGG 3499
Qy 3481 GCCAACCGCTTACCTGGAGGAGCGCAACATCTGTGACGCTCTGGGCTGACCACTCCTT 3540
Db 3500 GCCAACCGCTTACCTGGAGGAGCGCAACATCTGTGACGCTCTGGGCTGACCACTCCTT 3559
Qy 3541 GGCAGGATCTCTACTGAGTCCACCGCAGCAGAGATGATCGAGCGTGTGGAGAAAGC 3600
Db 3560 GGCAGGATCTCTACTGAGTCCACCGCAGCAGAGATGATCGAGCGTGTGGAGAAAGC 3619
Qy 3601 ACCGGGACAGCGGACTCGCATCCAGGCGCTGTGCGCCACCTCCTGAGCATCCATGCA 3660
Db 3620 ACCGGGACAGCGGACTCGCATCCAGGCGCTGTGCGCCACCTCCTGAGCATCCATGCA 3679
Qy 3661 GTGAGGAGAGTCAAGCTGGAGGAGTCTTACGCCCAACCATGTGCCCTGCAATGGTGGC 3720
Db 3680 GTGAGGAGAGTCAAGCTGGAGGAGTCTTACGCCCAACCATGTGCCCTGCAATGGTGGC 3739
Qy 3721 TGCTCCCACTCTGTATTTGCAAGGGTGTGGGACACACCGGTGTCTATGCGGCTGCAAC 3780
Db 3740 TGCTCCCACTCTGTATTTGCAAGGGTGTGGGACACACCGGTGTCTATGCGGCTGCAAC 3799
Qy 3781 CTCGTGCTCTGCAAGACCTGCTGACCTGTGAGAGCGGCGCCACTGCTCCCGGACGAG 3840
Db 3800 CTCGTGCTCTGCAAGACCTGCTGACCTGTGAGAGCGGCGCCACTGCTCCCGGACGAG 3859
Qy 3841 TTTTCATGTGCAACAGGGGAGATCGATCTATCCCGGGGCTTGGCGCTGTGAGCGGCTTT 3900
Db 3860 TTTTCATGTGCAACAGGGGAGATCGATCTATCCCGGGGCTTGGCGCTGTGAGCGGCTTT 3919
Qy 3901 CCCGAGTGGATGACAGAGCGAGAGGGCTTGGCCCTGTGCTTCCCGCCCGCAGTTC 3960
Db 3920 CCCGAGTGGATGACAGAGCGAGAGGGCTTGGCCCTGTGCTTCCCGCCCGCAGTTC 3979

500 GTGCTCTTCTGCGAGGACTTTGACACAGCCGAGGCGCATCGCTTGGACCCCGCTCACGGG 559
541 TACATGTACTGACAGACTGGGGTGAGACGCCCGCGGATTGAGCGGGCAGGGATGATGGC 600
560 TACATGTACTGACAGACTGGGGTTGAGACGCCCGCGGATTGAGCGGGCAGGGATGATGGC 619
601 AGCACCGGAGAGATCATTTGTGGACTCGGACATTTACTGGCCCAATGGACTGACCATCGAC 660
620 AGCACCGGAGAGATCATTTGTGGACTCGGACATTTACTGGCCCAATGGACTGACCATCGAC 679
661 CTGAGGAGCAGAGAAGCTTACTTGGGCTGACGCCAAGCTCAGCTTTCATCCACCGTGCCAAAC 720
680 CTGAGGAGCAGAGAAGCTTACTTGGGCTGACGCCAAGCTCAGCTTTCATCCACCGTGCCAAAC 739
721 CTGAGCGGCTGTTTCCGCGACAGAGTGTGGAGGGCAGCTGACGCCACCCCTTGCCTCG 780
740 CTGAGCGGCTGTTTCCGCGACAGAGTGTGGAGGGCAGCTGACGCCACCCCTTGCCTCG 799
781 ACGCTCTCCGGGACACTCTGTACTTGGACAGACTGGCAGACCCCGCTCCATCCATGCCCTGC 840
800 ACGCTCTCCGGGACACTCTGTATCTGGACAGACTGGCAGACCCCGCTCCATCCATGCCCTGC 859
841 AACAAAGCGCACTGGGGGGAAGAGGAGATCTCTGAGTGCCCTCTACTCAACCCATGGAC 900
860 AACAAAGCGCACTGGGGGGAAGAGGAGATCTCTGAGTGCCCTCTACTCAACCCATGGAC 919
901 ATCCAGGTGTGAGCAGGAGCGGAGCGCTTCTTTCACACTCTCGTGTGAGGAGACAAAT 960
920 ATCCAGGTGTGAGCAGGAGCGGAGCGCTTCTTTCACACTCTCGTGTGAGGAGACAAAT 979
961 GGGCGCTGCCACCTGTGCTGTCTGCCCAAGCGAGCGCTTCTACACATGCGCCTGC 1020
980 GGGCGCTGCCACCTGTGCTGTCTGCCCAAGCGAGCGCTTCTACACATGCGCCTGC 1039
1021 CCCACGGGTGTGAGCTGCAGGACAAACGGCAGGACGTGTAAAGGAGGAGCGCGAGGAGGTG 1080
1040 CCCACGGGTGTGAGCTGCAGGACAAACGGCAGGACGTGTAAAGGAGGAGCGCGAGGAGGTG 1099
1081 CTGCTGTGCGCGCGGCGAGCGGACCTTACGAGAGATCTCGCTGGACACGCCGAGCTTCAAC 1140
1100 CTGCTGTGCGCGCGGCGAGCGGACCTTACGAGAGATCTCGCTGGACACGCCGAGCTTCAAC 1159
1141 GACATCGTGTGAGCTGCAGGACGACATCCGCGACGCCATTTGCCATCGACTAGCAGCCGCTA 1200
1160 GACATCGTGTGAGCTGCAGGAGGACGACATCCGCGACGCCATTTGCCATCGACTAGCAGCCGCTA 1219
1201 GAGGCTATGTCTACTGGACAGATGACAGAGGTGGCGGCCATCCCGAGGCGGTACTCTGGAC 1260
1220 GAGGCTATGTCTACTGGACAGATGACAGAGGTGGCGGCCATCCCGAGGCGGTACTCTGGAC 1279
1261 GGGTCTGGGGGCGAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGTC 1320
1280 GGGTCTGGGGGCGAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGTC 1339
1321 GACTGGGTGGCCGAAACCTCTACTGGACCGACGCGGACCGGACCGCATCGAGGTGACG 1380
1340 GACTGGGTGGCCGAAACCTCTACTGGACCGACGCGGACCGGACCGCATCGAGGTGACG 1399
1381 GCGCTCAACGGGACCTCCCGCAAGATCTGTGTGTCGAGGAGACCTGGACGAGCCCGGAGCC 1440
1400 GCGCTCAACGGGACCTCCCGCAAGATCTGTGTGTCGAGGAGACCTGGACGAGCCCGGAGCC 1459
1441 ATCGCACTGACACCGCGTATGGGCTCATGTACTGACAGACTGGGGAGAGAACCTTAAA 1500
1460 ATCGCACTGACACCGCGTATGGGCTCATGTACTGACAGACTGGGGAGAGAACCTTAAA 1519
1501 ATCGAGTGTCCAACTTGGATGGGCGAGGCGGCTGTGCTGCTCAATGCGCTCCCTCGG 1560
1520 ATCGAGTGTCCAACTTGGATGGGCGAGGCGGCTGTGCTGCTCAATGCGCTCCCTCGG 1579
1561 TGGCCCAACGGGCTGGCCCTTGACCTCGCAGAGGGGGAAGCTCTACTTGGGGAGACGCCAAG 1620

1580 TGGCCCAACGGGCTGGCCCTTGGACCTGACGAGGAGGGAAGCTCTACTTGGGGAGACGCCAAG 1639
1621 ACAGACAGATCGAGGTGATCAATGTTGATGGGACGAGAGGCGGACCTCTCTCGAGGAC 1680
1640 ACAGACAGATCGAGGTGATCAATGTTGATGGGACGAGAGGCGGACCTCTCTCGAGGAC 1699
1681 AAGCTCCCGCACATTTTTCGGGTTTTCACGCTGCTGGGGGACTTTCATCTACTGAGTACTGG 1740
1700 AAGCTCCCGCACATTTTTCGGGTTTTCACGCTGCTGGGGGACTTTCATCTACTGAGTACTGG 1759
1741 CAGCGCGCAGCATCGAGCGGGTGCAACAAGTCAAGGCGAGCGGAGCGTCAATATTGAC 1800
1760 CAGCGCGCAGCATCGAGCGGGTGCAACAAGTCAAGGCGAGCGGAGCGTCAATATTGAC 1819
1801 CAGCTGCCCGACCTGATGGGGCTCAAAAGCTGTGAAATGTGGCCAAAGGCTGTGGAGACCAAC 1860
1820 CAGCTGCCCGACCTGATGGGGCTCAAAAGCTGTGAAATGTGGCCAAAGGCTGTGGAGACCAAC 1879
1861 CCGTGTGCGGACAGGAAACGGGGGTGAGCCACCTGTGCTTCTTTCACACCCCGCAAC 1920
1880 CCGTGTGCGGACAGGAAACGGGGGTGAGCCACCTGTGCTTCTTTCACACCCCGCAAC 1939
1921 CCGTGTGCGTGGCCCATCGGCTGAGAGTGTGATGACATGAAGAAGCTGTGATCGTGCCT 1980
1940 CCGTGTGCGTGGCCCATCGGCTGAGAGTGTGATGACATGAAGAAGCTGTGATCGTGCCT 1999
1981 GAGGCTTCTTGGTCTTTCACACGAGAGCGGCCATCCACAGGATCTCCCTCGAGACCAAT 2040
2000 GAGGCTTCTTGGTCTTTCACACGAGAGCGGCCATCCACAGGATCTCCCTCGAGACCAAT 2059
2041 AACAAACGAGCTGGGCATCCCGCTCAAGGGGTCAAGGAGGCTCAGGCCCTGGACTTTGAT 2100
2060 AACAAACGAGCTGGGCATCCCGCTCAAGGGGTCAAGGAGGCTCAGGCCCTGGACTTTGAT 2119
2101 GTGTCCAAACCAATCTACTTGAACAGAGTGTGAGCTGTGAAGACCAATCAGCGCGGCTTC 2160
2120 GTGTCCAAACCAATCTACTTGAACAGAGTGTGAGCTGTGAAGACCAATCAGCGCGGCTTC 2179
2161 ATGAAACGGGAGCTCGGTGGAGACAGTGTGAGTGTGGCTTGTACTACCCCGAGGGCATG 2220
2180 ATGAAACGGGAGCTCGGTGGAGACAGTGTGAGTGTGGCTTGTACTACCCCGAGGGCATG 2239
2221 GCCCTTGAATGGATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2280
2240 GCCCTTGAATGGATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2299
2281 GTGCGCGGCTGAGACGGGCAAGTTCGGCAAGTCTCGTGTGGAGGAGCTTGGACCAACCG 2340
2300 GTGCGCGGCTGAGACGGGCAAGTTCGGCAAGTCTCGTGTGGAGGAGCTTGGACCAACCG 2359
2341 AGGTCTGCTGGCCCTGGATCCCAAGGGCTTACATCTACTGAGCGGAGTGGGGCGGCAAG 2400
2360 AGGTCTGCTGGCCCTGGATCCCAAGGGCTTACATCTACTGAGCGGAGTGGGGCGGCAAG 2419
2401 CCGAGGATCTGTGCGGGCTTTCATGGAACGGGACCAACTGATGACGCTGGTGGACAGGCTG 2460
2420 CCGAGGATCTGTGCGGGCTTTCATGGAACGGGACCAACTGATGACGCTGGTGGACAGGCTG 2479
2461 GGGCGGGCGCAACGAGCTTCCATTTGACTTACGCTGACAGCGGCTTACTTGGACCGGCTG 2520
2480 GGGCGGGCGCAACGAGCTTCCATTTGACTTACGCTGACAGCGGCTTACTTGGACCGGCTG 2539
2521 GACACCAACATGATCGGATCGTCCAAATGTGCTGGGTGAGGCGGCTGTGATTTGCCGAC 2580
2540 GACACCAACATGATCGGATCGTCCAAATGTGCTGGGTGAGGCGGCTGTGATTTGCCGAC 2599
2581 GATCTCTCCGCAACCGGCTTCCGCTTTCAGCGAGTATGATGATGATGATGATGATGATGATGAT 2640
2600 GATCTCTCCGCAACCGGCTTCCGCTTTCAGCGAGTATGATGATGATGATGATGATGATGATGAT 2659
2641 AATCTGCAAGCATTTGAGCGGGCGGACAAAGATGAGCGGCGGAAACCGCACCTCATCCAG 2700
2660 AATCTGCAAGCATTTGAGCGGGCGGACAAAGATGAGCGGCGGAAACCGCACCTCATCCAG 2719

||||| 1489 ATCGACTGCACCCCGTGTATGGGCTCATGTACTGACAGACTGGGGAGAGAACCTTAA 1548
QY ||||| 1501 ATCGAGTGTGCAACTTGGATGGGAGGAGCGCGTGTGCTGGTCAATGCCCTCCCTCGGG 1560
Db ||||| 1549 ATCGAGTGTGCAACTTGGATGGGAGGAGCGCGTGTGCTGGTCAATGCCCTCCCTCGGG 1608
QY ||||| 1561 TGGCCCAACGGCTGGCCCTGACCTGACGAGGGGAGAGCTTACTTGGGGAGAGCCCAAG 1620
Db ||||| 1609 TGGCCCAACGGCTGGCCCTGACCTGACGAGGGGAGAGCTTACTTGGGGAGAGCCCAAG 1668
QY ||||| 1621 ACAGACAAGATCGAGGTGATCAATGTGATGAGGACGAAAGAGCGGACCTCTCTGAGGAC 1680
Db ||||| 1669 ACAGACAAGATCGAGGTGATCAATGTGATGAGGACGAAAGAGCGGACCTCTCTGAGGAC 1728
QY ||||| 1681 AAGCTCCGCGCAATTTTGGGTTTCACTGCTGGGGGACTTCACTACTGGAAGTACTGG 1740
Db ||||| 1729 AAGCTCCGCGCAATTTTGGGTTTCACTGCTGGGGGACTTCACTACTGGAAGTACTGG 1788
QY ||||| 1741 CAGCGCCGAGCATCGAGGGGTGCAAGGTCAAGGTCAAGGCGAGCGGAGCGTCAATTTGAC 1800
Db ||||| 1789 CAGCGCCGAGCATCGAGGGGTGCAAGGTCAAGGTCAAGGCGAGCGGAGCGTCAATTTGAC 1848
QY ||||| 1801 CAGCTGCCCGCACTGATGGGGCTCAAAGCTGTGAATGTGGCCAAAGGTCTGCGAACCAAC 1860
Db ||||| 1849 CAGCTGCCCGCACTGATGGGGCTCAAAGCTGTGAATGTGGCCAAAGGTCTGCGAACCAAC 1908
QY ||||| 1861 CGGTGTGGGACAGAAACGGGGGTGACAGCCACTGTGCTTTTCAACCCCAAGCAAC 1920
Db ||||| 1909 CGGTGTGGGACAGAAACGGGGGTGACAGCCACTGTGCTTTTCAACCCCAAGCAAC 1968
QY ||||| 1921 CGGTGTGGCTGCCCATCGGCTGTGAGCTGTGAGTGAATGAGACATGAGACCTGTGCT 1980
Db ||||| 1969 CGGTGTGGCTGCCCATCGGCTGTGAGCTGTGAGTGAATGAGACATGAGACCTGTGCT 2028
QY ||||| 1981 GAGGCTTCTTGGTCTTCAACAGACAGCGCGCATCCACAGATCTCCCTCGAGACCAAT 2040
Db ||||| 2029 GAGGCTTCTTGGTCTTCAACAGACAGCGCGCATCCACAGATCTCCCTCGAGACCAAT 2088
QY ||||| 2041 AACAAAGAGTGGCCATCCCGTCAACGGGGTCAAGGAGGCTCAGCCCTGGACTTTGAT 2100
Db ||||| 2089 AACAAAGAGTGGCCATCCCGTCAACGGGGTCAAGGAGGCTCAGCCCTGGACTTTGAT 2148
QY ||||| 2101 GTGTCCAAACCAACATCTACTGACAGACCTCAGCTGAGACCATCAGCGCGGCTTC 2160
Db ||||| 2149 GTGTCCAAACCAACATCTACTGACAGACCTCAGCTGAGACCATCAGCGCGGCTTC 2208
QY ||||| 2161 ATGAAACGGGAGCTCGGTGGAGCAGTGTGGAGTTTGGCTTTGACTACCCCGAGGGCATG 2220
Db ||||| 2209 ATGAAACGGGAGCTCGGTGGAGCAGTGTGGAGTTTGGCTTTGACTACCCCGAGGGCATG 2268
QY ||||| 2221 GCGGTTGATGGGCAAGAACTTCTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2280
Db ||||| 2269 GCGGTTGATGGGCAAGAACTTCTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2328
QY ||||| 2281 GTGGCGGGCTGAGCGGCGAGTTCGGCAAGTCTCGGTGGAGGAGCTTGGACAAACCG 2340
Db ||||| 2329 GTGGCGGGCTGAGCGGCGAGTTCGGCAAGTCTCGGTGGAGGAGCTTGGACAAACCG 2388
QY ||||| 2341 AGGTGCTGGCCCTGGATCCCAACAAAGGCTTACATCTACTGGACCGAGTGGGGGGCAAG 2400
Db ||||| 2389 AGGTGCTGGCCCTGGATCCCAACAAAGGCTTACATCTACTGGACCGAGTGGGGGGCAAG 2448
QY ||||| 2401 CCGAGGATCGGGGCTTCTATGAGCGGAGCAACTGATGACGCTGGTGGACAGGTG 2460
Db ||||| 2449 CCGAGGATCGGGGCTTCTATGAGCGGAGCAACTGATGACGCTGGTGGACAGGTG 2508
QY ||||| 2461 GCGCGGGCAACGACCTCACCATTGACTTACCTGACCAAGCGGCTTCTACTGACCGACTG 2520
Db ||||| 2509 GCGCGGGCAACGACCTCACCATTGACTTACCTGACCAAGCGGCTTCTACTGACCGACTG 2568
QY ||||| 2521 GACACCAACATGATCGAGTCTGCAACATGCTGGGTGAGGACGGGTGATTTGCCGAC 2580

Db ||||| 2569 GACACCAACATGATCGAGTCTGCTCMAACATGCTGGGTGAGGAGCGGGTCTGTGATTCGCGAC 2628
QY ||||| 2581 GATCTCCGCGCACCCGTTCTGGTCTTGACGAGTACAGCGATTATATCTACTTGACAGACTGG 2640
Db ||||| 2629 GATCTCCGCGCACCCGTTCTGGTCTTGACGAGTACAGCGATTATATCTACTTGACAGACTGG 2688
QY ||||| 2641 AATCTGACAGATTTAGCGGGCGGACAAAGACTAGCGGGCGGAAACCGCACCTCTATCCAG 2700
Db ||||| 2689 AATCTGACAGATTTAGCGGGCGGACAAAGACTAGCGGGCGGAAACCGCACCTCTATCCAG 2748
QY ||||| 2701 GGCACCTTGGACTTCTGTGATGAGACATCTCTGTGTTCACCTCTCCGCGAGGATGGCTTC 2760
Db ||||| 2749 GGCACCTTGGACTTCTGTGATGAGACATCTCTGTGTTCACCTCTCCGCGAGGATGGCTTC 2808
QY ||||| 2761 AATGACTGTATGCAACAACCGGCGAGTGTGGGAGCTGTGCTTGGCCATCCCCGGCGGC 2820
Db ||||| 2809 AATGACTGTATGCAACAACCGGCGAGTGTGGGAGCTGTGCTTGGCCATCCCCGGCGGC 2868
QY ||||| 2821 CACGCTCGGGCTGGGCTTCACTACACCTTGACCCAGAGCGGAGCGGACCTGACGCGCG 2880
Db ||||| 2869 CACGCTCGGGCTGGGCTTCACTACACCTTGACCCAGAGCGGAGCGGACCTGACGCGCG 2928
QY ||||| 2881 CCCACCACTTCTGTGCTTTCAGCCAGAAATCTGCATCAGTTCGATGATCCCGGAGCGAC 2940
Db ||||| 2929 CCCACCACTTCTGTGCTTTCAGCCAGAAATCTGCATCAGTTCGATGATCCCGGAGCGAC 2988
QY ||||| 2941 CAGCACAGCCCGGATCTCATCTCTCCCTGATGAGTGAAGAAACGTCAAGACCATCGAC 3000
Db ||||| 2989 CAGCACAGCCCGGATCTCATCTCTCCCTGATGAGTGAAGAAACGTCAAGACCATCGAC 3048
QY ||||| 3001 TATGACCCACTGGAACAAGTTTCATCTACTGGGTGATGGGCGCCAGAACATCAAGCGAGCC 3060
Db ||||| 3049 TATGACCCACTGGAACAAGTTTCATCTACTGGGTGATGGGCGCCAGAACATCAAGCGAGCC 3108
QY ||||| 3061 AAGCACAGCGGACCCAGGCTTCTGTGCTTGTGACTCTCTGAGCCAGGCGCAAAACCCAGAC 3120
Db ||||| 3109 AAGCACAGCGGACCCAGGCTTCTGTGCTTGTGACTCTCTGAGCCAGGCGCAAAACCCAGAC 3168
QY ||||| 3121 AGGCAAGCCCAAGACTCTGAGCATCGACATCTPAACGCGGACACTGTGCTTGAAGCTGCGAG 3180
Db ||||| 3169 AGGCAAGCCCAAGACTCTGAGCATCGACATCTPAACGCGGACACTGTGCTTGAAGCTGCGAG 3228
QY ||||| 3181 GCCACCAATACCAATCAAGCTCCAGAGCTGAGCGGGGAGCCATGGGGGTGGTCTGGT 3240
Db ||||| 3229 GCCACCAATACCAATCAAGCTCCAGAGCTGAGCGGGGAGCCATGGGGGTGGTCTGGT 3288
QY ||||| 3241 GGGGACCGGACCAAGCCAGGCTTCTGTCTAAACGCGAGCGAGGCTTACTGTACTTC 3300
Db ||||| 3289 GGGGACCGGACCAAGCCAGGCTTCTGTCTAAACGCGAGCGAGGCTTACTGTACTTC 3348
QY ||||| 3301 ACCAAATGCAAGGACCGGGACGCAAGATCGAAGCGGAGCGCTTGGACCGGACCGAGCGC 3360
Db ||||| 3349 ACCAAATGCAAGGACCGGGACGCAAGATCGAAGCGGAGCGCTTGGACCGGACCGAGCGC 3408
QY ||||| 3361 GAGTCTCTTTCACACCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
Db ||||| 3409 GAGTCTCTTTCACACCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3468
QY ||||| 3421 GGCAGAGCTGTTCTGGGTGAGACCGGAGCTTGAAGCGCATTTGAGAGCTGTGACTGTGAGGG 3480
Db ||||| 3469 GGCAGAGCTGTTCTGGGTGAGACCGGAGCTTGAAGCGCATTTGAGAGCTGTGACTGTGAGGG 3528
QY ||||| 3481 GGCAGAGCTGTTCTGGGTGAGAGCGGCAACATGCTGAGCGCTTGGGCTGAGACCATCTT 3540
Db ||||| 3529 GGCAGAGCTGTTCTGGGTGAGAGCGGCAACATGCTGAGCGCTTGGGCTGAGACCATCTT 3588
QY ||||| 3541 GGCAGAGCTTCTGACTGATCGAGCGGACCGGACGAGATGATCGAGCGCTTGGAGAGGAC 3600
Db ||||| 3589 GGCAGAGCTTCTGACTGATCGAGCGGACCGGACGAGATGATCGAGCGCTTGGAGAGGAC 3648
QY ||||| 3601 ACCGGGACAAAGGAGCTTCGATCCAGGGCGGTGCTGCCCACTCTACTTGGCATCCATGCA 3660
Db ||||| 3649 ACCGGGACAAAGGAGCTTCGATCCAGGGCGGTGCTGCCCACTCTACTTGGCATCCATGCA 3708

Db 229 CGGCTGGTGGAGCCCGCGGAGTCAAGCTGGAGTCCACCATCGTGGTCAAGCGGCTGGAG 288
QY 241 GATCGCGCGCAGTGGACTTCCAGTCTTCCAAAGGAGCGCGTCTACTTGGACAGACCTGGAGC 300
Db 289 GATCGCGCGCAGTGGACTTCCAGTCTTCCAAAGGAGCGCGTCTACTTGGACAGACCTGGAGC 348
QY 301 GAGGAGGCGCATCAAGCAGACCTACTGAAACAGACGCGGCGCGCGTGGAGAACCTGGTTC 360
Db 349 GAGGAGGCGCATCAAGCAGACCTACTGAAACAGACGCGGCGCGCGTGGAGAACCTGGTTC 408
QY 361 ATCTCCGGCTCGTCTCTCCGACGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 420
Db 409 ATCTCCGGCTCGTCTCTCCGACGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 468
QY 421 TGGACGGACTCAGACACCAACCGCATCGAGTGGCCAACTCAATGGCAGCATCCCGGAAG 480
Db 469 TGGACGGACTCAGACACCAACCGCATCGAGTGGCCAACTCAATGGCAGCATCCCGGAAG 528
QY 481 GTGCTCTTCTGGCAGGACCTTGAACAGCTAGGCGCATCGCTTGGACCCCGCTCACGGG 540
Db 529 GTGCTCTTCTGGCAGGACCTTGAACAGCTAGGCGCATCGCTTGGACCCCGCTCACGGG 588
QY 541 TACATGTACTGGACAGACTGGGCTGAGACGCGCCCGATTTGAGCGGCGAGGATGATG3C 600
Db 589 TACATGTACTGGACAGACTGGGCTGAGACGCGCCCGATTTGAGCGGCGAGGATGATG3C 648
QY 601 AGCACCCGGAAGATCAITGTGGACTCGGACATTTACTGGCCCAATGGACTGACCATCGAC 660
Db 649 AGCACCCGGAAGATCAITGTGGACTCGGACATTTACTGGCCCAATGGACTGACCATCGAC 708
QY 661 CTGGAGGAGCAGACCTCTACTGGCTGACGCGCAGCTCAGCTTCACTCAGCGTGCAC 720
Db 709 CTGGAGGAGCAGACCTCTACTGGCTGACGCGCAGCTCAGCTTCACTCAGCGTGCAC 768
QY 721 CTGGAGGAGCAGACCTCTCGGAGAGTGGTGGAGGCGACCTGACGACCCCTTGGCCCTG 780
Db 769 CTGGAGGAGCAGACCTCTCGGAGAGTGGTGGAGGCGACCTGACGACCCCTTGGCCCTG 828
QY 781 AGCTCTCTGGGAGACACTCTGTATGGAAGAGTGGGAGAGCCCGTCCATCGATGCTGC 840
Db 829 AGCTCTCTGGGAGACACTCTGTATGGAAGAGTGGGAGAGCCCGTCCATCGATGCTGC 888
QY 841 AACAGCGCACTGGGGGAGAGGAGAGTCTGTAGTGGCTCTACTCACCACATGGAC 900
Db 889 AACAGCGCACTGGGGGAGAGGAGAGTCTGTAGTGGCTCTACTCACCACATGGAC 948
QY 901 ATCCAGGTGCTGAGCCAGGAGCGGAGCCTTTCTTCCACACTCGCTGTGAGGAGCAAT 960
Db 949 ATCCAGGTGCTGAGCCAGGAGCGGAGCCTTTCTTCCACACTCGCTGTGAGGAGCAAT 1008
QY 961 GCGGCTGCTCCACCTGTGCTGTCTGTCCTCCAGCGAGCCTTCTACATAGCGCTGC 1020
Db 1009 GCGGCTGCTCCACCTGTGCTGTCTGTCCTCCAGCGAGCCTTCTACATAGCGCTGC 1068
QY 1021 CCCACGGGTGTCAGCTGCAGACAAACGCGAGGAGCTGTAGGAGGAGGAGGAGGAGG 1080
Db 1069 CCCACGGGTGTCAGCTGCAGACAAACGCGAGGAGCTGTAGGAGGAGGAGGAGGAGG 1128
QY 1081 CTGCTGTGCGCGCGGAGCGGACCTACGAGGATCTCGCTGGGACGCGGAGCTTCAAC 1140
Db 1129 CTGCTGTGCGCGCGGAGCGGACCTACGAGGATCTCGCTGGGACGCGGAGCTTCAAC 1188
QY 1141 GACATCGTGTGAGGTGGAGCAGCATCCGACCGCCATTGCGCATGAGCAGCCGCTA 1200
Db 1189 GACATCGTGTGAGGTGGAGCAGCATCCGACCGCCATTGCGCATGAGCAGCCGCTA 1248
QY 1201 GAGGCTATGTCTACTGACAGATGACGAGGTGGGCGCATCCGACGCGGCTACTGAC 1260
Db 1249 GAGGCTATGTCTACTGACAGATGACGAGGTGGGCGCATCCGACGCGGCTACTGAC 1308
QY 1261 GGGTCTGGGGCGAGACGCTGGTCAACCGAGATCAACGACCCCGATGGCATCGCGTTC 1320

Db 1309 GGGTCTGGGGCGAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCCGCTC 1368
QY 1321 GACTGGGTGGCCCGAAACCTCTACTTGGACCCACACCGGACCGCATCGAGTTCAGC 1380
Db 1369 GACTGGGTGGCCCGAAACCTCTACTTGGACCCACACCGGACCGCATCGAGTTCAGC 1428
QY 1381 CGCTTCAACGGGACCTCCGCAAGATCTGTGTGTCGAGGACCTTGGACGAGCCCGAGCC 1440
Db 1429 CGCTTCAACGGGACCTCCGCAAGATCTGTGTGTCGAGGACCTTGGACGAGCCCGAGCC 1488
QY 1441 ATCGCACTGCAACCGCTGATGGGCTCATGTACTTGGACAGACTGGGAGAGAACCTTAA 1500
Db 1489 ATCGCACTGCAACCGCTGATGGGCTCATGTACTTGGACAGACTGGGAGAGAACCTTAA 1548
QY 1501 ATCGAGTGTGCAACTTGGATGGGCGAGGAGCGGCTGTGTGTGTCATCTCCCTCCG 1560
Db 1549 ATCGAGTGTGCAACTTGGATGGGCGAGGAGCGGCTGTGTGTGTCATCTCCCTCCG 1608
QY 1561 TGGCCCAACCGCTTGGCCCTTGGACCTTGACAGAGGAGGAGCTTACTTGGGAGAGCCCAAG 1620
Db 1609 TGGCCCAACCGCTTGGCCCTTGGACCTTGACAGAGGAGGAGCTTACTTGGGAGAGCCCAAG 1668
QY 1621 ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAGGAGGAGCTTACTTGGGAGAGCCCAAG 1680
Db 1669 ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAGGAGGAGCTTACTTGGGAGAGCCCAAG 1728
QY 1681 AAGCTCCGCAATTTTGGGTTTCACTGTCTGGGGGACTTCTACTTGGGAGTGAATGG 1740
Db 1729 AAGCTCCGCAATTTTGGGTTTCACTGTCTGGGGGACTTCTACTTGGGAGTGAATGG 1788
QY 1741 CAGCGCGCAGCATCGAGCGGTCGACAGGTGCAAGGTCGAGCGCGGAGCTCATCTTGCAC 1800
Db 1789 CAGCGCGCAGCATCGAGCGGTCGACAGGTGCAAGGTCGAGCGCGGAGCTCATCTTGCAC 1848
QY 1801 CAGCTGCCGACCTGATGGGCTCAAAAGCTGTGAATGTGGCCAAAGGTGCTCGGAAACCAAC 1860
Db 1849 CAGCTGCCGACCTGATGGGCTCAAAAGCTGTGAATGTGGCCAAAGGTGCTCGGAAACCAAC 1908
QY 1861 CCGTGTGGGACAGAAACGCGGGGTGAGCCACTGTGTCTTCTTCAACCCCAAGCAAC 1920
Db 1909 CCGTGTGGGACAGAAACGCGGGGTGAGCCACTGTGTCTTCTTCAACCCCAAGCAAC 1968
QY 1921 CCGTGTGGTCCCGCCATCGCTGAGTGTGTGATGACATGAAGACCTGATGCTGCTCCT 1980
Db 1969 CCGTGTGGTCCCGCCATCGCTGAGTGTGTGATGACATGAAGACCTGATGCTGCTCCT 2028
QY 1981 GAGGCTCTTGTGTCTTCAACGAGCAGGCGGCCATCCACAGGATCTCCCTCGAGACCAAT 2040
Db 2029 GAGGCTCTTGTGTCTTCAACGAGCAGGCGGCCATCCACAGGATCTCCCTCGAGACCAAT 2088
QY 2041 AACAAACGAGTGGCCATCCGCTCACCGGCGTCAAGGAGGCTCAGCGCTGGACTTTGAT 2100
Db 2089 AACAAACGAGTGGCCATCCGCTCACCGGCGTCAAGGAGGCTCAGCGCTGGACTTTGAT 2148
QY 2101 GTGTCCAAACCAACATCTACTTGGACAGAGCTCAGCTGAAAGACCATCAGCGCGCTTC 2160
Db 2149 GTGTCCAAACCAACATCTACTTGGACAGAGCTCAGCTGAAAGACCATCAGCGCGCTTC 2208
QY 2161 ATGAAACGCGGAGCTCGTGGAGCAGCTGTGGAGTTTGGCTTGAATACCCCGAGGCGATG 2220
Db 2209 ATGAAACGCGGAGCTCGTGGAGCAGCTGTGGAGTTTGGCTTGAATACCCCGAGGCGATG 2268
QY 2221 GCGCTTCACTGGATGGGCAAGAACCTCTACTGGGCGGACCTGGGACCAACAGATCGAA 2280
Db 2269 GCGCTTCACTGGATGGGCAAGAACCTCTACTGGGCGGACCTGGGACCAACAGATCGAA 2328
QY 2281 GTGCGCGGCTGGACGCGGCAAGTTCCCGCAAGTCTCCTGTGGAGGAGCTTGGACCAACCG 2340
Db 2329 GTGCGCGGCTGGACGCGGCAAGTTCCCGCAAGTCTCCTGTGGAGGAGCTTGGACCAACCG 2388
QY 2341 AGTGTGCTGGCCCTGGATCCCAACGAGGCTTACATCTACTTGGACCGAGTGGGGCGGAG 2400
Db 2389 AGTGTGCTGGCCCTGGATCCCAACGAGGCTTACATCTACTTGGACCGAGTGGGGCGGAG 2448

QY 2401 CCGAGGATCGTGGCGGCTTCAATGACCGGACCAACTGCATGACGCTGGTGGACAAGGTG 2460
DB 2449 CCGAGGATCGTGGCGGCTTCAATGACCGGACCAACTGCATGACGCTGGTGGACAAGGTG 2508
QY 2461 GCGCGGGCAACGACCTCACATTTGACTACGCTGACAGCGCTCTACTGACGACGACTG 2520
DB 2509 GCGCGGGCAACGACCTCACATTTGACTACGCTGACAGCGCTCTACTGACGACGACTG 2568
QY 2521 GACACCAACATGATCGAGTCTGCAACATGCTGGGTGAGGAGCGGCTGCTATTTGCCGAC 2580
DB 2569 GACACCAACATGATCGAGTCTGCAACATGCTGGGTGAGGAGCGGCTGCTATTTGCCGAC 2628
QY 2581 GATCTCCGCGCACCGCTTGGGTCTGACGAGTACGAGCAATATATCTACTGACAGACTGG 2640
DB 2629 GATCTCCGCGCACCGCTTGGGTCTGACGAGTACGAGCAATATATCTACTGACAGACTGG 2688
QY 2641 AATCTGCACAGATTGAGCGGGCGGCAAGACTAGCGCGCGGAAACCGCACCTCATCCAG 2700
DB 2689 AATCTGCACAGATTGAGCGGGCGGCAAGACTAGCGCGCGGAAACCGCACCTCATCCAG 2748
QY 2701 GGCACCTGGACTTCTGATGAGATCTCTGGGTGTTCCACTCTCCCGCAGGATGGCTC 2760
DB 2749 GGCACCTGGACTTCTGATGAGATCTCTGGGTGTTCCACTCTCCCGCAGGATGGCTC 2808
QY 2761 AATGACTGTATGCAACAAACGCGGAGTGTGGGAGCTGTGCTTGCCTCCCGCGGCGC 2820
DB 2809 AATGACTGTATGCAACAAACGCGGAGTGTGGGAGCTGTGCTTGCCTCCCGCGGCGC 2868
QY 2821 CACCGCTGGCGTGGCTCTCACTACACCTGTGACCCGACGCGCAATGACGCGCG 2880
DB 2869 CACCGCTGGCGTGGCTCTCACTACACCTGTGACCCGACGCGCAATGACGCGCGC 2928
QY 2881 CCCACCACTTCTGCTGTGTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGGAGC 2940
DB 2929 CCCACCACTTCTGCTGTGTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGGAGC 2988
QY 2941 CAGCACAGCCCGGATCTCATCTGCTCCCTGCATGGACTGAGGAACTGCAAGCCATCGAC 3000
DB 2989 CAGCACAGCCCGGATCTCATCTGCTCCCTGCATGGACTGAGGAACTGCAAGCCATCGAC 3048
QY 3001 TATGACCACTGAGCAAGTTTCACTACTGAGTGGATGGCGCCAGAAATCAAGCGAGCC 3060
DB 3049 TATGACCACTGAGCAAGTTTCACTACTGAGTGGATGGCGCCAGAAATCAAGCGAGCC 3108
QY 3061 AAGGACGAGCGGACCCAGCCCTTTGTTTGGACCTCTCTGAGCCGAGGCAAAACCCAGAC 3120
DB 3109 AAGGACGAGCGGACCCAGCCCTTTGTTTGGACCTCTCTGAGCCGAGGCAAAACCCAGAC 3168
QY 3121 AGGACGACCCGACCTCAGCATCGACATCTACAGCCGAGCACTGTTCTGAGCGTGGCGAG 3180
DB 3169 AGGACGACCCGACCTCAGCATCGACATCTACAGCCGAGCACTGTTCTGAGCGTGGCGAG 3228
QY 3181 GCCACCAATACCATCAACGTTCCACAGGCTGAGCGGGGAAGCCATGGGGGTGGTCTCGGT 3240
DB 3229 GCCACCAATACCATCAACGTTCCACAGGCTGAGCGGGGAAGCCATGGGGGTGGTCTCGGT 3288
QY 3241 GGGGACCGGACAGCCCGAGGCAATGCTGCTCAACCGGAGCGAGGGTACTGTTACTTTC 3300
DB 3289 GGGGACCGGACAGCCCGAGGCAATGCTGCTCAACCGGAGCGAGGGTACTGTTACTTTC 3348
QY 3301 ACCAACATGACGAGCCGCGGACCAAGATCGAAACGCGCAGCCCTGGAGCGGACCGAGCGC 3360
DB 3349 ACCAACATGACGAGCCGCGGACCAAGATCGAAACGCGCAGCCCTGGAGCGGACCGAGCGC 3408
QY 3361 GAGGTCTCTTTCACACCGGCTCATCCGCCCTGTGGCCCTGGTGGTGGACAACACACTG 3420
DB 3409 GAGGTCTCTTTCACACCGGCTCATCCGCCCTGTGGCCCTGGTGGTGGACAACACACTG 3468
QY 3421 GGCAGCTGTTCTGGTGGACCGGACCTGAGCGCATTTGAGAGCTGTGACTGTGAGG 3480
DB 3469 GGCAGCTGTTCTGGTGGACCGGACCTGAGCGCATTTGAGAGCTGTGACTGTGAGG 3528

QY 3481 GCCAACCGCCTGACCTCTGGAGGACGCCAACATCGTGCAGCCTCTGGGCTGAGCATTCTT 3540
DB 3529 GCCAACCGCCTGACCTCTGGAGGACGCCAACATCGTGCAGCCTCTGGGCTGAGCATTCTT 3588
QY 3541 GGCAGGATCTCTACTGATTCGACCGCAGCAGACGATGATCGAGCGTGTGGAGAGACC 3600
DB 3589 GGCAGGATCTCTACTGATTCGACCGCAGCAGACGATGATCGAGCGTGTGGAGAGACC 3648
QY 3601 ACCGGGACAAAGCGGACTCGCATCCAGGCGGCTGTGCGCCACCTCAGTGGCATCCATGCA 3660
DB 3649 ACCGGGACAAAGCGGACTCGCATCCAGGCGGCTGTGCGCCACCTCAGTGGCATCCATGCA 3708
QY 3661 GTGAGGAAGTCACTCTGGAGGAGTCTTCAAGCCCAACCATGTGCGCGTGAACAATGGTGGC 3720
DB 3709 GTGAGGAAGTCACTCTGGAGGAGTCTTCAAGCCCAACCATGTGCGCGTGAACAATGGTGGC 3768
QY 3721 TGCTCCACATCTGTATTTGCCAAGGGTGTATGGGACACCAACGGTGTCTATGCGCCAGTCCAC 3780
DB 3769 TGCTCCACATCTGTATTTGCCAAGGGTGTATGGGACACCAACGGTGTCTATGCGCCAGTCCAC 3828
QY 3781 CTCGTGCTCTGCAAGAACCTGCTGACCTGTGAGAGCGCGCCACCTGCTCCCGGACCGAG 3840
DB 3829 CTCGTGCTCTGCAAGAACCTGCTGACCTGTGAGAGCGCGCCACCTGCTCCCGGACCGAG 3888
QY 3841 TTTGCAATGTGCAACAGGGGAGATCGACTGTATCCCGGGGCTGTGGCGTGTGAACGGCTTT 3900
DB 3889 TTTGCAATGTGCAACAGGGGAGATCGACTGTATCCCGGGGCTGTGGCGTGTGAACGGCTTT 3948
QY 3901 CCCGAGTGCATGATGACAGAGCGAGAGGGCTGTGCGCGTGTGGTGTGCGCGCGCCAGTTC 3960
DB 3949 CCCGAGTGCATGATGACAGAGCGAGAGGGCTGTGCGCGTGTGGTGTGCGCGCGCCAGTTC 4008
QY 3961 CCTGCGCGGGGGTCACTGTGTGAGACTGCTGCGCTGCGCTGCGAGCGGAGGAGACTGT 4020
DB 4009 CCTGCGCGGGGGTCACTGTGTGAGACTGCTGCGCTGCGCTGCGAGCGGAGGAGACTGT 4068
QY 4021 CAGGACCGCTCAGACGAGGGGAGTGTGAGCCCATCTGCTTCCCAACACGATTCGGGTGT 4080
DB 4069 CAGGACCGCTCAGACGAGGGGAGTGTGAGCCCATCTGCTTCCCAACACGATTCGGGTGT 4128
QY 4081 GCGAGCGGCTGATGTGCTCTCATCAACAGAGTGTGAGTCTTCTTCCCGGACTGTATCGAC 4140
DB 4129 GCGAGCGGCTGATGTGCTCTCATCAACAGAGTGTGAGTCTTCTTCCCGGACTGTATCGAC 4188
QY 4141 GGCTCCGACGAGCTCATGTGAAATCACCAAGCGGCTCAGACGACGAGCGCGGCCAC 4200
DB 4189 GGCTCCGACGAGCTCATGTGAAATCACCAAGCGGCTCAGACGACGAGCGCGGCCAC 4248
QY 4201 AGCAGTGCCATCGGGGCCCTCATTTGGCATCATCTCTCTCTTCTGTCATGGGTGGTGTG 4260
DB 4249 AGCAGTGCCATCGGGGCCCTCATTTGGCATCATCTCTCTCTCTCTCTCTCTCTCTCTCT 4308
QY 4261 TATTTTGTGTGTCAGCGCGTGTGTCAGGCTATGCGGGGGGCAACGCGGCCCTTCCCG 4320
DB 4309 TATTTTGTGTGTCAGCGCGTGTGTCAGGCTATGCGGGGGGCAACGCGGCCCTTCCCG 4368
QY 4321 CACGAGTATGTACGCGGAGCCCGCAGCTGCTCAATTTTCATAGCCCGGGGGTTC 4380
DB 4369 CACGAGTATGTACGCGGAGCCCGCAGCTGCTCAATTTTCATAGCCCGGGGGTTC 4428
QY 4381 CAGCATGGCCCTTTCACAGGCTCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCCTG 4440
DB 4429 CAGCATGGCCCTTTCACAGGCTCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCCTG 4488
QY 4441 ATGGGGGCGCGGGGCGGGTGTCCCTGTACGACGAGAAACAGTCAAGGGGCTCTGTTC 4500
DB 4489 ATGGGGGCGCGGGGCGGGTGTCCCTGTACGACGAGAAACAGTCAAGGGGCTCTGTTC 4548
QY 4501 AGCAGCTCTGTCAGACGAGGAGCGGCTGTACCGCGGATCTGAAACCGCGCCCTTC 4560
DB 4549 AGCAGCTCTGTCAGACGAGGAGCGGCTGTACCGCGGATCTGAAACCGCGCCCTTC 4608
QY 4561 CCGGCCACGAGCCCTCTCTGTACAAACATGAGCATGTTCTTCTTCAAAACATTCGGGCC 4620

[illegible]

Search completed: February 19, 2005, 06:21:24
Job time : 2609 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2005, 01:15:53 / Search time 201 seconds
(without alignments)
3107.552 Million cell updates/sec

Title: US-09-931-375A-2

Perfect score: 8740

Sequence: 1 MEAAPGPPWPLLLLLLLLLL.....TERSVFHLPPPPSPCTDSS 1615

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8740	100.0	1615	5	Aae21730 Human bon
2	8740	100.0	1615	6	Ab411131 Human LRP
3	8740	100.0	1615	7	Ab598798 Human Zma
4	8740	100.0	1615	8	Ab084659 Human can
5	8736	100.0	1615	2	Aaw83309 LRP5 prot
6	8736	100.0	1615	5	Aae21740 Human BSM
7	8736	100.0	1615	8	Adi27181 Human LRP
8	8736	100.0	1615	8	Adq20524 Human sof
9	8736	100.0	1615	8	Ad17560 Human hig
10	8736	100.0	1615	8	Ad48211 Human hig
11	8736	100.0	1615	8	Ad73482 Human low
12	8735	99.9	1615	5	Aae21741 Human BSM
13	8727	99.9	1615	8	Ad17561 Human hig
14	8727	99.9	1615	8	Ad48212 Human hig
15	8724	99.8	1627	7	Bsm85665 Human pro
16	8724	99.8	1627	8	Ab084660 Human can
17	8715	99.7	1615	4	Aag68169 Human Zma
18	8715	99.7	1615	6	Ab411093 Human wil
19	8715	99.7	1615	7	Ab598058 Human LRP
20	8715	99.7	1615	7	Ad682427 Human Zma
21	8715	99.7	1615	8	Ad16921 Human hig
22	8715	99.7	1615	8	Ad47572 Human hig
23	8706	99.6	1615	4	Aag68170 Human HBM
24	8706	99.6	1615	5	Aau80880 Human hig
25	8706	99.6	1615	6	Ab411094 Human LRP

26	8706	99.6	1615	7	ADB98059	LRP5 mute
27	8706	99.6	1615	7	ADe82428	Human HBM
28	8706	99.6	1615	8	Ad16922	Human hig
29	8706	99.6	1615	8	ADR47573	Human hig
30	8705	99.6	1611	8	AD127180	Human LRP
31	8705	99.6	1665	6	ABR41133	Human LRP
32	8705	99.6	1665	7	ADB98800	Human LRP
33	8702	99.6	1615	5	AAU80879	Human Zma
34	8607	98.5	1591	2	AAW83308	Mature LRP
35	8574	98.1	1639	2	AAW83311	LRP5 isoform
36	8259.5	94.5	1614	2	AAW83312	Mouse LRP
37	8259.5	94.5	1614	5	ABW07255	Mouse LRP
38	8259.5	94.5	1614	8	AD127179	Mouse LRP
39	8254.5	94.4	1614	8	AD127193	Mouse LRP
40	8254.5	94.4	1614	8	AD127174	Mouse LRP
41	7972.5	91.2	1564	6	ABR41132	Mouse LRP
42	7972.5	91.2	1564	7	ADB98799	Mouse Zma
43	7882	90.2	1451	2	AAW83310	LRP5 prot
44	6868	78.6	1325	8	ABO84658	Mouse can
45	6153.5	70.4	1613	6	ABR41134	Human LRP

ALIGNMENTS

RESULT 1
AAE21730
ID AAE21730 standard; protein; 1615 AA.

XX AC AAE21730;

XX 16-JUL-2002 (first entry)

XX Human bone strength and mineralisation regulatory protein (BSMR).

XX Human; bone strength and mineralisation regulatory protein; BSMR;

XX bone strength; mineralisation; ophthalmological; antidiabetic;

XX bone density regulating transmembrane receptor; prosthetic device;

XX surgical implant; diabetic retinopathy; hypertensive retinopathy;

XX therapy; osteoporosis; prematurity; ocular vessel; eye disorder;

XX osteopathic.

XX Homo sapiens.

XX WO200216553-A2.

XX 28-FEB-2002.

XX 17-AUG-2001; 2001WO-US041788.

XX 18-AUG-2000; 2000US-0226119P.

XX 22-SEP-2000; 2000US-0234337P.

XX 13-JUL-2001; 2001US-0304851P.

XX (AVET) AVENTIS PHARMA SA.

XX (HARD) HARVARD COLLEGE.

XX (UYCA-) UNIV CASE WESTERN RESERVE.

XX Warman ML, Gong Y, Olsen BR, Rawadi G, Roman-Roman S;

XX WPI; 2002-329694/36.

XX N-PSDB; AAD34322.

XX Polynucleotide encoding bone strength and mineralization regulatory protein useful for diagnosis or therapy of osteoporosis.

XX Disclosure; Fig 3; 124pp; English.

XX The invention relates to bone strength and mineralisation regulatory protein (BSMR) and its corresponding nucleic acid sequence. BSMR DNA is useful for the diagnosis or therapy of osteoporosis and for regulating (increasing) bone strength and mineralisation in a human subject by activating a bone density regulating transmembrane receptor (BSMR)

CC protein). An expression vector comprising a promoter that is operably
 CC linked to BSMR DNA is useful for modulating bone density and for
 CC enhancing bone strength and mineralisation in a mammal cell. Composition
 CC comprising a BSMR effector is useful for treating osteoporosis and is
 CC useful particularly as a coating for prosthetic devices and surgical
 CC implants. BSMR is useful for screening lead pharmaceutical agents as BSMR
 CC effectors, which may be used to treat a range of eye disorders such as
 CC diabetic retinopathy, hypertensive retinopathy and retinopathy of
 CC prematurity, in which normal vascular growth and integrity of ocular
 CC vessels is disrupted. The present sequence is human BSMR protein
 XX
 SQ Sequence 1615 AA;

Query Match	100.0%;	Score 8740;	DB 5;	Length 1615;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1615;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	MEAAAPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRVLVDAGGVKLESTIVVS	60
DB	1	MEAAAPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRVLVDAGGVKLESTIVVS	60
QY	61	GLEDAAAVDQPSKGAIVYTDVSEBAIKQTYLNQTAAGVQNVVLSGLVSPDGLACDWVGK	120
DB	61	GLEDAAAVDQPSKGAIVYTDVSEBAIKQTYLNQTAAGVQNVVLSGLVSPDGLACDWVGK	120
QY	121	KLYTDSSTNRLEVANLNGTSRKVLFWQDLDOPRAIALDPAGYNYWTDWGETPRIERAG	180
DB	121	KLYTDSSTNRLEVANLNGTSRKVLFWQDLDOPRAIALDPAGYNYWTDWGETPRIERAG	180
QY	181	MDGSTRKIIIVSDIYWPNGLTLDLEQKLYWADAKLSPIHRANLDGSPRQKVVEGSLTHP	240
DB	181	MDGSTRKIIIVSDIYWPNGLTLDLEQKLYWADAKLSPIHRANLDGSPRQKVVEGSLTHP	240
QY	241	FALTLSGDTLWTDQWTSIIHACNKRGTGKKEILSALYSPMDIQVLSEQRQPPFHTRCE	300
DB	241	FALTLSGDTLWTDQWTSIIHACNKRGTGKKEILSALYSPMDIQVLSEQRQPPFHTRCE	300
QY	301	EDNGGCSHLCLLSPSEPYTCACPTGVQDQNGRTCKAGAEVLLIARLTDLRLISLDT	360
DB	301	EDNGGCSHLCLLSPSEPYTCACPTGVQDQNGRTCKAGAEVLLIARLTDLRLISLDT	360
QY	361	DFTDVLQVDDIRHAIAIDYDPLEGYVYTDDEVRARIRAYLDGSGAQLVNTINDDPDG	420
DB	361	DFTDVLQVDDIRHAIAIDYDPLEGYVYTDDEVRARIRAYLDGSGAQLVNTINDDPDG	420
QY	421	IADVARNLWTDGTDRIEVTRNGTSRKILVSEDLDEPRALHHPVWGLMTWTDWGE	480
DB	421	IADVARNLWTDGTDRIEVTRNGTSRKILVSEDLDEPRALHHPVWGLMTWTDWGE	480
QY	481	NPKIECANLDQERRVLYNVAISLGNPGLALDLQEGKLYWDAKTDKIEVINVDGTKRRTL	540
DB	481	NPKIECANLDQERRVLYNVAISLGNPGLALDLQEGKLYWDAKTDKIEVINVDGTKRRTL	540
QY	541	LEDKLPPIFGTLTGDFIYTDWQRRSIRVHKVKAASRDVIIDQLPDLMLGLKAVNVAKV	600
DB	541	LEDKLPPIFGTLTGDFIYTDWQRRSIRVHKVKAASRDVIIDQLPDLMLGLKAVNVAKV	600
QY	601	GTNPCADRNGCSHLCPPTPHATRCGCPIGLELISDMKTCIVPEAFVFTSRAAIHRLS	660
DB	601	GTNPCADRNGCSHLCPPTPHATRCGCPIGLELISDMKTCIVPEAFVFTSRAAIHRLS	660
QY	661	ETNNNDVAIPTGVKEASALDFDVSNHIIYTDVSLKTIISRAFNGSGSVEHVEFGLDYP	720
DB	661	ETNNNDVAIPTGVKEASALDFDVSNHIIYTDVSLKTIISRAFNGSGSVEHVEFGLDYP	720
QY	721	EGMAVDMWNGKLYWADTGTNRIEVARLDQGRQVLVWRDLNPRSLALDPTKGYIYWTEW	780
DB	721	EGMAVDMWNGKLYWADTGTNRIEVARLDQGRQVLVWRDLNPRSLALDPTKGYIYWTEW	780
QY	781	GGKPRIVRAFDNGTNCMTLVKVGANDLTIDYADQRLYWTDLTNMIESNMILGQERV	840
DB	781	GGKPRIVRAFDNGTNCMTLVKVGANDLTIDYADQRLYWTDLTNMIESNMILGQERV	840

QY	841	IADDLPHPFGLTQYSDIYIYTDWNLHSHIERADKTSGRNRTLIQGHLDVFMOLLVPHSSRQ	900
DB	841	IADDLPHPFGLTQYSDIYIYTDWNLHSHIERADKTSGRNRTLIQGHLDVFMOLLVPHSSRQ	900
QY	901	DGLNDCMNNGCQQLCLAIIPGHRCCGASHYTLDPSSRNCSPPTTFLFSQKSAISRWI	960
DB	901	DGLNDCMNNGCQQLCLAIIPGHRCCGASHYTLDPSSRNCSPPTTFLFSQKSAISRWI	960
QY	961	PDDQHSPLLIPLHGLRNVAKAIDYDPLDKFYIYWDGRQNIKRAKDDGTQPPVLTSLSQGQ	1020
DB	961	PDDQHSPLLIPLHGLRNVAKAIDYDPLDKFYIYWDGRQNIKRAKDDGTQPPVLTSLSQGQ	1020
QY	1021	NPDRQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVNAERG	1080
DB	1021	NPDRQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVNAERG	1080
QY	1081	LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVVALVVDNTLGLFWVDADLKRISCD	1140
DB	1081	LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVVALVVDNTLGLFWVDADLKRISCD	1140
QY	1141	LSGANRLTLEDANIVQPLGLTILGKHLIYIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG	1200
DB	1141	LSGANRLTLEDANIVQPLGLTILGKHLIYIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG	1200
QY	1201	IHAVEEVSLEFSAHPCARDNGGCSHICIAKGDGTGPRCSPVHLVILONLLTTCGEPTCS	1260
DB	1201	IHAVEEVSLEFSAHPCARDNGGCSHICIAKGDGTGPRCSPVHLVILONLLTTCGEPTCS	1260
QY	1261	PDQFACATGEIDCIPGAWRCDGPPCEDDQSDDEGCPVCSAAQFFCARGQCVDLRLCDGE	1320
DB	1261	PDQFACATGEIDCIPGAWRCDGPPCEDDQSDDEGCPVCSAAQFFCARGQCVDLRLCDGE	1320
QY	1321	ADCDRDEADCDAILPNOFRCASGQCVLIIKQCDSPDCIDGSDLMCSITKPPSDDS	1380
DB	1321	ADCDRDEADCDAILPNOFRCASGQCVLIIKQCDSPDCIDGSDLMCSITKPPSDDS	1380
QY	1381	PAHSSAIGPVIGIITLSLFWMGVYFVQVVCQRYAGANGPPHYVSGTGHVPLNFTAP	1440
DB	1381	PAHSSAIGPVIGIITLSLFWMGVYFVQVVCQRYAGANGPPHYVSGTGHVPLNFTAP	1440
QY	1441	GGSGHGPFPTGIACGKSNMWSVSLMGGRGVPLVDRNHVTGASSSSSSSTKATLYPPIINP	1500
DB	1441	GGSGHGPFPTGIACGKSNMWSVSLMGGRGVPLVDRNHVTGASSSSSSSTKATLYPPIINP	1500
QY	1501	PPSPATDPSLYNMDFYSSNIPATARYPIIIRMGMAPPTTTCSTDVCDSDYSASRWKAS	1560
DB	1501	PPSPATDPSLYNMDFYSSNIPATARYPIIIRMGMAPPTTTCSTDVCDSDYSASRWKAS	1560
QY	1561	KYLDLNSDSDPYPPPTPHSQVLSAEDSCPPSPATERSYFHLPPPPSPCTDSS	1615
DB	1561	KYLDLNSDSDPYPPPTPHSQVLSAEDSCPPSPATERSYFHLPPPPSPCTDSS	1615

RESULT 2

ABR41131
 ID ABR41131 standard; protein; 1615 AA.

AC ABR41131;

XX
 DT 02-JUN-2003 (first entry)

XX Human LRP5 protein.

XX Human; high bone mass; HEM; LRP5; LRP6; transgenic; bone mass modulation;
 KW gene therapy; bone density modulation; bone strength; trabecular number;
 KW bone size; bone tissue connectivity; bone disease; osteoporosis;
 KW osteomalacia; rickets; Paget's disease; neoplasm of the bone.

OS Homo sapiens.

XX WO200292764-A2.

XX 21-NOV-2002.

PD

XX PF 13-MAY-2002; 2002WO-US014876.
XX PR 11-MAY-2001; 2001US-0290071P.
XX PR 17-MAY-2001; 2001US-0291311P.
XX PR 01-FEB-2002; 2002US-0353058P.
XX PR 04-MAR-2002; 2002US-0361293P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PA (AMHP) WYETH.
XX PI Babij P, Bex FJ, Yaworsky PJ, Bodine PV;
XX PI WPI; 2003-129278/12.
XX PT New transgenic animals (e.g. mice), useful as models for studying bone
XX PT density modulation, developing drugs for treating or preventing bone
XX PT diseases (e.g. osteoporosis), or diagnosing diseases characterized by
XX PT reduced bone density.
XX PS Disclosure; Fig 26; 603pp; English.
XX CC The invention relates to novel transgenic animals expressing the high
XX CC bone mass (HBM) gene, expressing the corresponding wild type HBM gene,
XX CC comprising an alteration of the gene encoding LRP5 or LRP6, or expressing
XX CC an LRP5 that is modulated by an altered gene control sequence introduced
XX CC by homologous or non-homologous recombination. The transgenic animals are
XX CC for the study of bone density modulation or bone mass modulation. The
XX CC invention has osteopathic and cytostatic activity. The polynucleotides of
XX CC the invention may have a use in gene therapy. The transgenic animals and
XX CC nucleic acids are for the study of bone density modulation, where the
XX CC bone mass is modulated relative to non-transgenic animals of the same
XX CC species in more than one parameter selected from bone density, bone
XX CC strength, trabecular number, bone size, or bone tissue connectivity. The
XX CC transgenic animals, nucleic acids and methods are useful for identifying
XX CC molecules involved in bone development, and for developing pharmaceutical
XX CC compositions, which may be employed for treating or preventing bone
XX CC diseases, e.g. osteoporosis, osteomalacia, rickets, Paget's disease, or
XX CC neoplasms of the bone. The transgenic animals and nucleic acids are also
XX CC useful in methods for diagnosing diseases involved in bone development, is
XX CC or characterised by reduced bone density or mass. The present sequence, is
XX CC used in the exemplification of the invention
SQ Sequence 1615 AA;
Query Match 100.0%; Score 8740; DB 6; Length 1615;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1615; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEAPPGPWPPLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVS 60
DB 1 MEAPPGPWPPLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVS 60
QY 61 GLEDAAAVDFQSKGAVYWTVDVSEAIKQTYLNTQTGAQVNVVIGSLVSPDGLACDWGK 120
DB 61 GLEDAAAVDFQSKGAVYWTVDVSEAIKQTYLNTQTGAQVNVVIGSLVSPDGLACDWGK 120
QY 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLQOPRAIALDPAGHYMYTWDWGETPRIERAG 180
DB 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLQOPRAIALDPAGHYMYTWDWGETPRIERAG 180
QY 181 MDGSTRKIIIVSDIYWPNGLIIDLEEQKLYWADAKLSFIHRANLDGSPKQKVVGSLTHP 240
DB 181 MDGSTRKIIIVSDIYWPNGLIIDLEEQKLYWADAKLSFIHRANLDGSPKQKVVGSLTHP 240
QY 241 FALTLSGDTLYWTDWQTRS IHACNKRGTGGRKEILSALYSPMDIQVLSQERQPFHTRCE 300
DB 241 FALTLSGDTLYWTDWQTRS IHACNKRGTGGRKEILSALYSPMDIQVLSQERQPFHTRCE 300
QY 301 EDNGGCSHLCLLSSEPPYTCACPTGVQLOQNGRTCKAGAEVLLAARTDLRLISLDTTP 360
DB 301 EDNGGCSHLCLLSSEPPYTCACPTGVQLOQNGRTCKAGAEVLLAARTDLRLISLDTTP 360

QY 361 DFTDIVLQVDDIRHAIADIDYDPLEGYVYVTTDDEVRAIRRAYLDGSGAQLVNTINDEPDG 420
DB 361 DFTDIVLQVDDIRHAIADIDYDPLEGYVYVTTDDEVRAIRRAYLDGSGAQLVNTINDEPDG 420
QY 421 IAVDWARNLYWTDGTGTDRIEIVTRINGTSRKILSVSEDLDEPRALHHPVGLMYWTDWGE 480
DB 421 IAVDWARNLYWTDGTGTDRIEIVTRINGTSRKILSVSEDLDEPRALHHPVGLMYWTDWGE 480
QY 481 NPKIECANLDGQERRVLVNASLIGWPNGLALDQSGKLYWGDAKTDKIEVINVDGTKRRTL 540
DB 481 NPKIECANLDGQERRVLVNASLIGWPNGLALDQSGKLYWGDAKTDKIEVINVDGTKRRTL 540
QY 541 LEDKLPHIFGFTLLGDFIYWTWQRRSIERVHKVKASRDVIIDQLPDLMLGLKAVNAVAVV 600
DB 541 LEDKLPHIFGFTLLGDFIYWTWQRRSIERVHKVKASRDVIIDQLPDLMLGLKAVNAVAVV 600
QY 601 GTNPCADRNGGCSHLCPFTPHATRCGPIGIELLSDMKTCIVPEAFVFTSRAAIIHRISL 660
DB 601 GTNPCADRNGGCSHLCPFTPHATRCGPIGIELLSDMKTCIVPEAFVFTSRAAIIHRISL 660
QY 661 ETNNNDVAIPUTGVKEASALDPDVSNHHIYWTDSLTKTISRFAFMNGSSVEHVVEFGLDYP 720
DB 661 ETNNNDVAIPUTGVKEASALDPDVSNHHIYWTDSLTKTISRFAFMNGSSVEHVVEFGLDYP 720
QY 721 EGMAVDWMGKNLYWADTGTNRIEVARLDGQPRQVLVWRDLNPRSLALDPTKGYIYTWTEW 780
DB 721 EGMAVDWMGKNLYWADTGTNRIEVARLDGQPRQVLVWRDLNPRSLALDPTKGYIYTWTEW 780
QY 781 GGGPRIIVRAFMGNTCMTLVDKVGGRANDLTIDYADQRLYWTDLTNMTIESSNMLGQSRV 840
DB 781 GGGPRIIVRAFMGNTCMTLVDKVGGRANDLTIDYADQRLYWTDLTNMTIESSNMLGQSRV 840
QY 841 IADLPHPGTGLTOYSDIYWTWNLHSTERADTKSGNRNLTIOGHLDPMVMDILVPHSSRQ 900
DB 841 IADLPHPGTGLTOYSDIYWTWNLHSTERADTKSGNRNLTIOGHLDPMVMDILVPHSSRQ 900
QY 901 DGLNDCMHNNGCCQLCLAIIPGHRGCGASHYTLDPSSRNCSPPTTFLFQSKAISRM 960
DB 901 DGLNDCMHNNGCCQLCLAIIPGHRGCGASHYTLDPSSRNCSPPTTFLFQSKAISRM 960
QY 961 PDDQHSPLILPLHGLRNKVIDYDPLDKFTYVWDGRQNIKRAKDDGTGTPVLTSLSGQ 1020
DB 961 PDDQHSPLILPLHGLRNKVIDYDPLDKFTYVWDGRQNIKRAKDDGTGTPVLTSLSGQ 1020
QY 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVHRLSSEAMGVVLRGDRDKPRAIVNAERG 1080
DB 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVHRLSSEAMGVVLRGDRDKPRAIVNAERG 1080
QY 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESC 1140
DB 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESC 1140
QY 1141 LSGANRLTLEDANIVQPLGLTILGKHLVWIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG 1200
DB 1141 LSGANRLTLEDANIVQPLGLTILGKHLVWIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG 1200
QY 1201 IHAVEEVSLEFSAPHCARDNGGCSHICIAKDGTPRCSVPVHLVQLNLITCCPEPTCS 1260
DB 1201 IHAVEEVSLEFSAPHCARDNGGCSHICIAKDGTPRCSVPVHLVQLNLITCCPEPTCS 1260
QY 1261 PDQFACATGEIDCIPGAWRCDGPFPECCDDQSDDEECPCVSAQAQPCARGQCVDLRLRCDGE 1320
DB 1261 PDQFACATGEIDCIPGAWRCDGPFPECCDDQSDDEECPCVSAQAQPCARGQCVDLRLRCDGE 1320
QY 1321 ADCQDRSDEADCAICLPNOFRCASGCVLIKQCDSPDPCIDGSDMLCMCITKPPSDS 1380
DB 1321 ADCQDRSDEADCAICLPNOFRCASGCVLIKQCDSPDPCIDGSDMLCMCITKPPSDS 1380
QY 1381 PAHSSAIPVIGIITLSLFWMGVYVFCORVVCORVYAGANGPPHVEYSGTGHVPLNFTAP 1440
DB 1381 PAHSSAIPVIGIITLSLFWMGVYVFCORVVCORVYAGANGPPHVEYSGTGHVPLNFTAP 1440
QY 1441 GGSQHGPFTGIACGKSMWSSVSLMGRGGVPLYDRNHVTGASSSSSSSTKATLYPPLNP 1500

```

Db 1441 GGSQAGPTGIACGKSMSSVSLMGGGVPYLDNRNHTVGASSSSSTKATLYPILNP 1500
QY 1501 PPSPATDSLNMDFYSSNIPATAPRYPYIIRGWAPPTTTCSTDVCDSDYSASRWKAS 1560
Db 1501 PPSPATDSLNMDFYSSNIPATAPRYPYIIRGWAPPTTTCSTDVCDSDYSASRWKAS 1560
QY 1561 KYLDLNSDSDPYPPPHPSOYLSAEDSCPSPATERSYFHLFPSPSPCTDSS 1615
Db 1561 KYLDLNSDSDPYPPPHPSOYLSAEDSCPSPATERSYFHLFPSPSPCTDSS 1615

RESULT 3
ID ADB98798
XX ADB98798 standard; protein; 1615 AA.
AC ADB98798;
DT 04-DEC-2003 (first entry)
DE Human Zmax1(LRP5).
XX Osteopathic; Gene therapy; High Bone Mass; HBM; LRP5; Zmax1; LRP6;
KW bone mass modulation; osteoporosis; human.
OS Homo sapiens.
XX WO200292000-A2.
PD 21-NOV-2002.
XX 13-MAY-2002; 2002WO-US014877.
PR 11-MAY-2001; 2001US-0290071P.
PR 17-MAY-2001; 2001US-0291311P.
PR 01-FEB-2002; 2002US-0353058P.
PR 04-MAR-2002; 2002US-0361293P.
XX (GENO-) GENOME THERAPEUTICS CORP.
PA (AMHP ) WYETH.
XX
PI Allen K, Anisowicz A, Graham JR, Morales A, Yaworsky PJ, Liu W;
DR WPI; 2003-129214/12.
XX
PT New nucleic acid comprising a mutation in LRP5 or LRP6, useful for
PT diagnosing a HBM-like phenotype in a subject and for preparing a
PT composition for modulating bone mass and/or lipid levels in a subject
PT suffering from e.g. osteoporosis.
XX
PS Disclosure; Fig 26; 629pp; English.
XX
CC The present invention relates to High Bone Mass (HBM), LRP5 (Zmax1) and
CC LRP6 mutants, which results in a HBM-like phenotype when expressed in a
CC cell. The HBM-like phenotype results in bone mass modulation and/or lipid
CC level modulation. The invention is useful for diagnosing a HBM-like
CC phenotype in a subject and for preparing a composition for modulating
CC bone mass and/or lipid levels in a subject suffering from e.g.
CC osteoporosis. The present sequence was used to illustrate the invention.
XX
SQ Sequence 1615 AA;

Query Match 100.0%; Score 8740; DB 7; Length 1615;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1615; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEAAPPGPPWPLLILLALLCGCPAPAAASPLLFFANRRDVRILVDAGGVKLESTIVVS 60
Db 1 MEAAPPGPPWPLLILLALLCGCPAPAAASPLLFFANRRDVRILVDAGGVKLESTIVVS 60
QY 61 GLEDAAAVDFOFSKGVVWTVDSBEAIKQTYLNQTAAGVQVNWISGLVSPDGLACDWVGK 120
Db 61 GLEDAAAVDFOFSKGVVWTVDSBEAIKQTYLNQTAAGVQVNWISGLVSPDGLACDWVGK 120

```

```

QY 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLDPRAIALDPAHGMYWTDWGETPRIERAG 180
Db 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLDPRAIALDPAHGMYWTDWGETPRIERAG 180
QY 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEOKLYWADAKLSPIHRANLDGSPROKVVESGLTHP 240
Db 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEOKLYWADAKLSPIHRANLDGSPROKVVESGLTHP 240
QY 241 FALTSLSGDTLYWTDQWTRSIHACNKRITGGKKEILSALYSMDIQVLSQERQPPPHTRCE 300
Db 241 FALTSLSGDTLYWTDQWTRSIHACNKRITGGKKEILSALYSMDIQVLSQERQPPPHTRCE 300
QY 301 EDNGCCHLCLLSPSEPPYTCACPTGYLOQNGTCKAGAEVILLAAARTDLRLISLTP 360
Db 301 EDNGCCHLCLLSPSEPPYTCACPTGYLOQNGTCKAGAEVILLAAARTDLRLISLTP 360
QY 361 DFTDVLQVDDIRHAIAIDYDPLEGYVYVWTDDEVRARRAYLDGSGAOTLVNTEINDDPG 420
Db 361 DFTDVLQVDDIRHAIAIDYDPLEGYVYVWTDDEVRARRAYLDGSGAOTLVNTEINDDPG 420
QY 421 IAVDWARNLYWTDGTDRIEVTRLNGTSRKILVSEDLDEPRALALHPVGLMYWTDWGE 480
Db 421 IAVDWARNLYWTDGTDRIEVTRLNGTSRKILVSEDLDEPRALALHPVGLMYWTDWGE 480
QY 481 NPKIECANLDGQERRVLVNASLGNLALDLQEGKLYWGDATKIEVINVDGTKRRTL 540
Db 481 NPKIECANLDGQERRVLVNASLGNLALDLQEGKLYWGDATKIEVINVDGTKRRTL 540
QY 541 LEDKLPHIFGFTLLGDFIYWTDWQRRSIERVHKYKASRDVIIDQLPDLMLKAVNAVAVV 600
Db 541 LEDKLPHIFGFTLLGDFIYWTDWQRRSIERVHKYKASRDVIIDQLPDLMLKAVNAVAVV 600
QY 601 GTNPCADRNGGCSHLCFFTHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRISL 660
Db 601 GTNPCADRNGGCSHLCFFTHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRISL 660
QY 661 ETNNNDVAIPLTVGKEASALDFVSNHHIYWTDVSLAKTISRAFMNGSSVEHVFGDLYP 720
Db 661 ETNNNDVAIPLTVGKEASALDFVSNHHIYWTDVSLAKTISRAFMNGSSVEHVFGDLYP 720
QY 721 EGMVDMWGNKLYWADTGTNRIEVARLDGQPROVLVWRDLNPRSLALDPTKGYIYWTEW 780
Db 721 EGMVDMWGNKLYWADTGTNRIEVARLDGQPROVLVWRDLNPRSLALDPTKGYIYWTEW 780
QY 781 GKGPRIVRAFMGDNVCMTLVDKVRANDLTIDYADQRLYWTDLDTNMIESNNMLGQSRVV 840
Db 781 GKGPRIVRAFMGDNVCMTLVDKVRANDLTIDYADQRLYWTDLDTNMIESNNMLGQSRVV 840
QY 841 IADDLPHFPGLTOYSDYIYWTDWNLHSTIERADKTSGRNRTLIOGHLDVMDILVFHSSRQ 900
Db 841 IADDLPHFPGLTOYSDYIYWTDWNLHSTIERADKTSGRNRTLIOGHLDVMDILVFHSSRQ 900
QY 901 DGLNDCMHNNGQCQOLCLAIIPGCHRCGCASHYTLDPSSRNCSPPTTFLFSQKSAISRM 960
Db 901 DGLNDCMHNNGQCQOLCLAIIPGCHRCGCASHYTLDPSSRNCSPPTTFLFSQKSAISRM 960
QY 961 PDDQHSDDLILPLHGLRNVAIDYDPLDKFIYVWVGQONIKRAKDDGTOPPVLTSLSGQ 1020
Db 961 PDDQHSDDLILPLHGLRNVAIDYDPLDKFIYVWVGQONIKRAKDDGTOPPVLTSLSGQ 1020
QY 1021 NPDRQPHDLSDIIVSRRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
Db 1021 NPDRQPHDLSDIIVSRRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
QY 1081 LYFTNMDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLAKRIESCD 1140
Db 1081 LYFTNMDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLAKRIESCD 1140
QY 1141 LSGANRLTLEDANIVQPLGLTILGKHLIYWDROQOMIERVEKTTGDKRTRIOGRAVHLTG 1200
Db 1141 LSGANRLTLEDANIVQPLGLTILGKHLIYWDROQOMIERVEKTTGDKRTRIOGRAVHLTG 1200

```


Db 541 LEDKLPHIFGTLGDFIYTWQRRSIEVHKVRSRVDIIDLQPDMLGLKANVAKV 600
QY 601 GNPACDRNGGSHLCFFPHATRCGPIGLLELLSDMTKTCIYPEAFLVFTSRAAHRISL 660
Db 601 GNPACDRNGGSHLCFFPHATRCGPIGLLELLSDMTKTCIYPEAFLVFTSRAAHRISL 660
QY 661 ETNNNDVAIPLTGVEASALDFVSNHHIYWTDSVLSKTIISRAFMGSSVEHVVEGLDYP 720
Db 661 ETNNNDVAIPLTGVEASALDFVSNHHIYWTDSVLSKTIISRAFMGSSVEHVVEGLDYP 720
QY 721 EGMAYDWMGKNLYADTGTNRLEVARLDGQFQVLRDLNPRSLALDPTKGYIYWTW 780
Db 721 EGMAYDWMGKNLYADTGTNRLEVARLDGQFQVLRDLNPRSLALDPTKGYIYWTW 780
QY 781 GGPRIIVAFMDGTNCMTLVKVGRLNDITIDYADQRLYWTDLDTNMIESSNMLGQERV 840
Db 781 GGPRIIVAFMDGTNCMTLVKVGRLNDITIDYADQRLYWTDLDTNMIESSNMLGQERV 840
QY 841 IADDLPHPEGLTOYSDYIYWTDLNLSIRADKTSGRNRTLIQHLDFVMDILVPHSSRQ 900
Db 841 IADDLPHPEGLTOYSDYIYWTDLNLSIRADKTSGRNRTLIQHLDFVMDILVPHSSRQ 900
QY 901 DGLNDCHMNGCGQLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLFSOKSAISMI 960
Db 901 DGLNDCHMNGCGQLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLFSOKSAISMI 960
QY 961 PDDQSPDLILPLHGLRNVAIDYDPLDKFIYWDGRQNIKRAKDDGTQPPVLTSLSGQ 1020
Db 961 PDDQSPDLILPLHGLRNVAIDYDPLDKFIYWDGRQNIKRAKDDGTQPPVLTSLSGQ 1020
QY 1021 NPDROPHDLSIDIYERTLFTWCEANTINVHLSGEAMGVLRGDRKPRALVNAERG 1080
Db 1021 NPDROPHDLSIDIYERTLFTWCEANTINVHLSGEAMGVLRGDRKPRALVNAERG 1080
QY 1081 LYFTNQDRAAKIERAALDGTREVLFTGLIRPVALVVDNTLGLFWVDADLKRIESCD 1140
Db 1081 LYFTNQDRAAKIERAALDGTREVLFTGLIRPVALVVDNTLGLFWVDADLKRIESCD 1140
QY 1141 LSGANRLTLEDANIYVQILGLTLGLHLYWIDRQQMIERVEKTDGKTRIQGRVAHLTG 1200
Db 1141 LSGANRLTLEDANIYVQILGLTLGLHLYWIDRQQMIERVEKTDGKTRIQGRVAHLTG 1200
QY 1201 IHAVEVSLEEPSAHPCARDNGGCSHICIAKDGTPRCSCPVHLVLLQNLITCGEPPTCS 1260
Db 1201 IHAVEVSLEEPSAHPCARDNGGCSHICIAKDGTPRCSCPVHLVLLQNLITCGEPPTCS 1260
QY 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEGCFVCSAAQFPFCARQCQVDLRLCDGE 1320
Db 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEGCFVCSAAQFPFCARQCQVDLRLCDGE 1320
QY 1321 ADCQDRSEADCDALCLPNQRCASQCQVLIKQCCDSDPDICDSDGDELMCEITPPSDS 1380
Db 1321 ADCQDRSEADCDALCLPNQRCASQCQVLIKQCCDSDPDICDSDGDELMCEITPPSDS 1380
QY 1381 PAHSSAIGPVTGIIILSLFVMGGVYFVQVRCORVAGANGPPPHYVSGTSPHVLNFTAP 1440
Db 1381 PAHSSAIGPVTGIIILSLFVMGGVYFVQVRCORVAGANGPPPHYVSGTSPHVLNFTAP 1440
QY 1441 GGSQHGPTGTIACGKSMSSVSLMGGRGGVPLYDRNHVTGASSSSSSSTKATLYPPIINP 1500
Db 1441 GGSQHGPTGTIACGKSMSSVSLMGGRGGVPLYDRNHVTGASSSSSSSTKATLYPPIINP 1500
QY 1501 PPSRATDSLNMDFYSSNIPATAPRYPYIIRGMAPPPTPCSDVCDSDYSASRWKAS 1560
Db 1501 PPSRATDSLNMDFYSSNIPATAPRYPYIIRGMAPPPTPCSDVCDSDYSASRWKAS 1560
QY 1561 KYIYDLNDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
Db 1561 KYIYDLNDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

AAW83309
ID AAW83309 standard; protein; 1615 AA.
AC AAW83309;
XX
DT 10-FEB-1999 (first entry)
XX
DE LRP5 protein from the longest open reading frame.
XX
KW LRP5; LDL-receptor related protein; LRP-3; IDDM; diabetes; endocytosis;
KW insulin dependent diabetes mellitus; autoimmune disease;
KW glomerulonephritis; inflammation; viral infection; osteoporosis;
KW hypercholesterolemia; Alzheimer's disease; low density lipoprotein.
XX
OS Homo sapiens.
XX
PN WO9846743-A1.
XX
PD 22-OCT-1998.
XX
PF 15-APR-1998; 98WO-GB0011102.
XX
PR 15-APR-1997; 97US-0043553P.
XX
PR 05-JUN-1997; 97US-0048740P.
XX
PA (WELL) WELLCOME TRUST LTD.
PA (MERI) MERCK & CO INC.
XX
XX
PI Todd JA, Hess JW, Caskey CT, Cox RD, Gerhold D, Hammond H;
PI Hey P, Kawaguchi Y, Merriman TR, Metzker ML, Nakagawa Y;
PI Phillips MS, Twells RCU;
XX
DR WPI; 1998-594573/50.
XX
PT New isolated LDL-receptor related protein - used to develop products for
PT treating, e.g. elevated triglyceride levels, diabetes, autoimmune
PT disorders, inflammation or Alzheimer's disease.
XX
PS Claim 2; Fig 5c; 200pp; English.
XX
The present sequence represents a LRP5 protein (low density lipoprotein
(LDL) receptor related protein, previously designated LRP-3). Nucleic
acid molecules (NAs) encoding LRP5 can be used for determining if an
individual is susceptible to insulin dependent diabetes mellitus (IDDM).
The NAs or proteins can be used for reducing triglyceride levels in the
serum of an individual. Therapies that affect LRP5 may also be useful in
the treatment of autoimmune diseases such as glomerulonephritis, diseases
and disorders involving disruption of endocytosis and/or antigen
presentation, cytokine clearance and/or inflammation, viral infection,
pathogenic bacterial toxin contamination, elevation of free fatty acids
or hypercholesterolemia, type 2 diabetes, osteoporosis, Alzheimer's
disease and cardiovascular disease. Products from the present invention
can also be used for detection, diagnosis and drug screening

Query Match 100.0%; Score 8736; DB 2; Length 1615;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MEAAPPPPPPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVLRLVAGGVKLESTIVS 60
Db 1 MEAAPPPPPPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVLRLVAGGVKLESTIVS 60
QY 61 GLEDAADVDFQSKGAVYWTDVSEAIKQYLNQGTGAQVNVVIGSLVSPDGLADWVGK 120
Db 61 GLEDAADVDFQSKGAVYWTDVSEAIKQYLNQGTGAQVNVVIGSLVSPDGLADWVGK 120
QY 121 KLYWTDSETRIEVANLNGTSRKVLFWQDLDPRAIALDPAHGYWYTDGCTPRIERAG 180
Db 121 KLYWTDSETRIEVANLNGTSRKVLFWQDLDPRAIALDPAHGYWYTDGCTPRIERAG 180
QY 181 MDGSTRKIIIVDSDIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPFRQKVEGSLTHP 240

Db 181 MDGSTRKIIIVDSDIYWPNGLTIDLEQKLYWADAKLSPIHRANLDGSRQKRVGSLTTP 240
QY 241 FALTILSGDTLYWTDWQTSIIHACNKRITGKGRKEILSALYSPMDIQVLQSRQPPFHTRCE 300
Db 241 FALTILSGDTLYWTDWQTSIIHACNKRITGKGRKEILSALYSPMDIQVLQSRQPPFHTRCE 300
QY 301 EDNGGCSHLCLLSPEPEYTCACPTGVOLQNGRTCKAGAEVLLIARRTDLRLISLDTLP 360
Db 301 EDNGGCSHLCLLSPEPEYTCACPTGVOLQNGRTCKAGAEVLLIARRTDLRLISLDTLP 360
QY 361 DFTDVLQVDDIRHAIADYDPLEGYVYWTDEVRRAIRRAYLDGSGAQTLVNTEINDPDG 420
Db 361 DFTDVLQVDDIRHAIADYDPLEGYVYWTDEVRRAIRRAYLDGSGAQTLVNTEINDPDG 420
QY 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRATALHPVWGLMYWTDWGE 480
Db 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRATALHPVWGLMYWTDWGE 480
QY 481 NPKIECANLDQERRVLYNVAISLWPNGLALDLQEGKLYWDAKTDKIEVINVDGTKRRTL 540
Db 481 NPKIECANLDQERRVLYNVAISLWPNGLALDLQEGKLYWDAKTDKIEVINVDGTKRRTL 540
QY 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIEVHKVKSARDVIIDQLPDLMLKAVNAVAVV 600
Db 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIEVHKVKSARDVIIDQLPDLMLKAVNAVAVV 600
QY 601 GTNPCADRNGCSHLCPFTPHATRCGCPIGIELISDMKTCIVPEAFVFTSRAAIIHRLS 660
Db 601 GTNPCADRNGCSHLCPFTPHATRCGCPIGIELISDMKTCIVPEAFVFTSRAAIIHRLS 660
QY 661 ETNNNDVAIPLTGVKEASALDPVSNHHIYWDVSLKTIISAFWNGSSVEHVFEGLDYP 720
Db 661 ETNNNDVAIPLTGVKEASALDPVSNHHIYWDVSLKTIISAFWNGSSVEHVFEGLDYP 720
QY 721 EGMVAVDMGKGLYNWADTGTNRIEVARLDGQFQVLRWDLNPRSLALDPTKGIYIYTEW 780
Db 721 EGMVAVDMGKGLYNWADTGTNRIEVARLDGQFQVLRWDLNPRSLALDPTKGIYIYTEW 780
QY 781 GKGPRIVRAFDGNTCNMTLVKVGRENDLTIDYADQRLYWTDLTNMISSNMLQGERVV 840
Db 781 GKGPRIVRAFDGNTCNMTLVKVGRENDLTIDYADQRLYWTDLTNMISSNMLQGERVV 840
QY 841 IADDLPHFPGLTQYSDIYWTDMNLHSIERADKTSGRNRTLIQGHLDFFWMDILVPHFSRQ 900
Db 841 IADDLPHFPGLTQYSDIYWTDMNLHSIERADKTSGRNRTLIQGHLDFFWMDILVPHFSRQ 900
QY 901 DGLNDCMHNNGCQGLCLAI PGHRCGCASHYTLDPSSRNCSPPTTFLFQSKAISRM 960
Db 901 DGLNDCMHNNGCQGLCLAI PGHRCGCASHYTLDPSSRNCSPPTTFLFQSKAISRM 960
QY 961 PDDQHSPLIILPLHGLRNVAIDYDPLDKFYWVDGRQNIKRAKDGTGTPFLVLSQGG 1020
Db 961 PDDQHSPLIILPLHGLRNVAIDYDPLDKFYWVDGRQNIKRAKDGTGTPFLVLSQGG 1020
QY 1021 NPDRQPHDLSIDIYSRITLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVNAERG 1080
Db 1021 NPDRQPHDLSIDIYSRITLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVNAERG 1080
QY 1081 LYFTNQMDRAAKIERAALDGTGEREVLTFTGILIRPVVALVVDNTLGLKFWVDADLKRIE 1140
Db 1081 LYFTNQMDRAAKIERAALDGTGEREVLTFTGILIRPVVALVVDNTLGLKFWVDADLKRIE 1140
QY 1141 LSGANRLTLEANIYVQPLGLTILGKHLWIDRQOQMIERVEKTTGDKRTRIOGRVAHLTG 1200
Db 1141 LSGANRLTLEANIYVQPLGLTILGKHLWIDRQOQMIERVEKTTGDKRTRIOGRVAHLTG 1200
QY 1201 IHAVEEVSLEFSAHPCARDNGGCSHCIAKGDGTPRCSVPHLVLLQNLLTTCGEPPTCS 1260
Db 1201 IHAVEEVSLEFSAHPCARDNGGCSHCIAKGDGTPRCSVPHLVLLQNLLTTCGEPPTCS 1260
QY 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEGCPVCSAAQFFPCARGQCVDLRLRCDGE 1320

Db 1361 PDQFACATGEIDCIPGAWRCDFPECDQSDDEGCPVCSAAQFFPCARGQCVDLRLRCDGE 1320
QY 1321 ADCQDRSDEADCAICLPNQRCASGQCVLIKQOCDSPDCIDGSDLMCMCITKPPSDDS 1380
Db 1321 ADCQDRSDEADCAICLPNQRCASGQCVLIKQOCDSPDCIDGSDLMCMCITKPPSDDS 1380
QY 1381 PAHSSAIGPVTIGIILSLFVWGGVYFVQVVCQRYAGANGFPPEHYVSGTTPHVLNFIAP 1440
Db 1381 PAHSSAIGPVTIGIILSLFVWGGVYFVQVVCQRYAGANGFPPEHYVSGTTPHVLNFIAP 1440
QY 1441 GGSQHGFPTGIACGKSMSSVSLMGGRGVPLYDRNHVHTGASSSSSSSTKATLYPILNP 1500
Db 1441 GGSQHGFPTGIACGKSMSSVSLMGGRGVPLYDRNHVHTGASSSSSSSTKATLYPILNP 1500
QY 1501 PPSPATPSLYNMDMFYSSNIPATRPVRYIIRGMAPPPTPCSTDVDCSDYSASRWKAS 1560
Db 1501 PPSPATPSLYNMDMFYSSNIPATRPVRYIIRGMAPPPTPCSTDVDCSDYSASRWKAS 1560
QY 1561 KYLLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLFPPPPSPCTDSS 1615
Db 1561 KYLLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLFPPPPSPCTDSS 1615

RESULT 6

AAE21740

ID AAE21740 standard; protein; 1615 AA.

XX AAE21740;

XX 16-JUL-2002 (first entry)

XX Human BSMR protein mutant, R494Q.

XX Human; bone strength and mineralisation regulatory protein; BSMR;
bone strength; mineralisation; ophthalmological; antidiabetic;
bone density regulating transmembrane receptor; prosthetic device;
surgical implant; diabetic retinopathy; hypertensive retinopathy;
therapy; osteoporosis; prematurity; ocular vessel; eye disorder;
osteopathic; mutant; mutain.

OS Homo sapiens.

XX Synthetic.

XX Key Location/Qualifiers

XX Misc-difference 494

XX /note= "Wild-type Arg substituted with Gln"

XX WO200216553-A2.

XX 28-FEB-2002.

XX 17-AUG-2001; 2001WO-US041788.

XX 18-AUG-2000; 2000US-0226119P.

XX 22-SEP-2000; 2000US-0234337P.

XX 13-JUL-2001; 2001US-0304851P.

XX (AVET) AVENTIS PHARMA SA.

XX (HARD) HARVARD COLLEGE.

XX (UYCA-) UNIV CASE WESTERN RESERVE.

XX Warman ML, Gong Y, Olsen BR, Rawadi G, Roman-Roman S;

XX WPI; 2002-329694/36.

XX Polynucleotide encoding bone strength and mineralization regulatory

XX protein useful for diagnosis or therapy of osteoporosis.

XX Disclosure; Page; 124pp; English.

XX The invention relates to bone strength and mineralisation regulatory

XX protein (BSMR) and its corresponding nucleic acid sequence. BSMR DNA is

XX useful for the diagnosis or therapy of osteoporosis and for regulating

(increasing) bone strength and mineralisation in a human subject by activating a bone density regulating transmembrane receptor (BSMR protein). An expression vector comprising a promoter that is operably linked to BSMR DNA is useful for modulating bone density and for enhancing bone strength and mineralisation in a mammal cell. Composition comprising a BSMR effector is useful for treating osteoporosis and is useful particularly as a coating for prosthetic devices and surgical implants. BSMR is useful for screening lead pharmaceutical agents as BSMR effectors, which may be used to treat a range of eye disorders such as diabetic retinopathy, hypertensive retinopathy and retinopathy of prematurity, in which normal vascular growth and integrity of ocular vessels is disrupted. The present sequence is human BSMR protein mutant. Note: This sequence is not shown in the specification, however it is constructed based on the human BSMR protein shown in fig 3 of the specification (AAE21730)

XX SQ Sequence 1615 AA;

Query Match 100.0%; Score 8736; DB 5; Length 1615;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1614; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEAAPGPPWPLLILLILLALCGCPAPAAAPLILLFANRRDRLVLDAGGVKLESTIVVS 60
 DB 1 MEAAPGPPWPLLILLILLALCGCPAPAAAPLILLFANRRDRLVLDAGGVKLESTIVVS 60

QY 61 GLEDAADVDFQSKAVYWDVSEAIKQTYLNQTAAGVONVVISGLVSPDGLACDWGK 120
 DB 61 GLEDAADVDFQSKAVYWDVSEAIKQTYLNQTAAGVONVVISGLVSPDGLACDWGK 120

QY 121 KLYWTDSETNRLEVANLNGTSRKVLFWQDLOPRAIALDPAGHYMYWTDWGETPRIERAG 180
 DB 121 KLYWTDSETNRLEVANLNGTSRKVLFWQDLOPRAIALDPAGHYMYWTDWGETPRIERAG 180

QY 181 MDGSTRKIIVDSDIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPRQKVVEGSLTHP 240
 DB 181 MDGSTRKIIVDSDIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPRQKVVEGSLTHP 240

QY 241 FALTISGDTLYWTDQWTSIHACNKRGTGKRKEILSALYSPMDIQVLQSERQPPFHTRCE 300
 DB 241 FALTISGDTLYWTDQWTSIHACNKRGTGKRKEILSALYSPMDIQVLQSERQPPFHTRCE 300

QY 301 EDNGGCSHLCLLSPSEPFVTCAPTGVOLQDNGRTCKAGAEVILLARRTLRLRLSLDTP 360
 DB 301 EDNGGCSHLCLLSPSEPFVTCAPTGVOLQDNGRTCKAGAEVILLARRTLRLRLSLDTP 360

QY 361 DPTDIVLQVDDIRHAIAIDYDPLEGYVYWDDEVRAIRRAVLDGSGAQLVNTENDPDG 420
 DB 361 DPTDIVLQVDDIRHAIAIDYDPLEGYVYWDDEVRAIRRAVLDGSGAQLVNTENDPDG 420

QY 421 IAVDVARNLYWTDGTDRIEVTRLNGTSRKILVSEDLDEPRAIALHPVWGLMYWTDWGE 480
 DB 421 IAVDVARNLYWTDGTDRIEVTRLNGTSRKILVSEDLDEPRAIALHPVWGLMYWTDWGE 480

QY 481 NPKIECANLDGQERRVLVNASLWENGLALDQEGKLYWGAQTKDIEVINVDGTKRRTL 540
 DB 481 NPKIECANLDGQERRVLVNASLWENGLALDQEGKLYWGAQTKDIEVINVDGTKRRTL 540

QY 541 LEDKLPHIFGFTLLGDFIYTDWQRRSIERVHKVKSARDVIIDQLPDLMLGLKAVNAVAVV 600
 DB 541 LEDKLPHIFGFTLLGDFIYTDWQRRSIERVHKVKSARDVIIDQLPDLMLGLKAVNAVAVV 600

QY 601 GTNPCADNRGGCSHLCPFTPHATRCGCPIGLELLSDMKTICVPEAFVFTSRAIHRISL 660
 DB 601 GTNPCADNRGGCSHLCPFTPHATRCGCPIGLELLSDMKTICVPEAFVFTSRAIHRISL 660

QY 661 ETNNNDVAIPTGVKEASALDFDVSNNHITYWTDVSLKTIISRAFMNGSSVEHWFEGLDYP 720
 DB 661 ETNNNDVAIPTGVKEASALDFDVSNNHITYWTDVSLKTIISRAFMNGSSVEHWFEGLDYP 720

QY 721 EGMAYDWMGKNLYWADTGTNRIEVARLDGQFQVLRWDLNPRSLALDPTKGYIYWTEW 780
 DB 721 EGMAYDWMGKNLYWADTGTNRIEVARLDGQFQVLRWDLNPRSLALDPTKGYIYWTEW 780

QY 781 GSKPRIVAFMGVNCMTLVKVGANDLTIDYADQRLYTDLDTNMTIENSNMLQSERV 840
 DB 781 GSKPRIVAFMGVNCMTLVKVGANDLTIDYADQRLYTDLDTNMTIENSNMLQSERV 840

QY 841 IADDLPHFPGLTQYSDYIYWDNLHSTERADKTSGRNRTLIOGHLDVWMDILFHSRQ 900
 DB 841 IADDLPHFPGLTQYSDYIYWDNLHSTERADKTSGRNRTLIOGHLDVWMDILFHSRQ 900

QY 901 DGLNDCMHNNCGGQOLCIAIPGHRCCGASHYTLDPSSRNCSPPTTFLLFQKSAISMI 960
 DB 901 DGLNDCMHNNCGGQOLCIAIPGHRCCGASHYTLDPSSRNCSPPTTFLLFQKSAISMI 960

QY 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFIYWDGRQNIKRAKDDGTQFPVLTSLSQGG 1020
 DB 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFIYWDGRQNIKRAKDDGTQFPVLTSLSQGG 1020

QY 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRKPRAIIVNAERG 1080
 DB 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRKPRAIIVNAERG 1080

QY 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESC 1140
 DB 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESC 1140

QY 1141 LSGANRLTLEDANIVQPLGLTILGKHLIYIDROQOMIERVEKTTGDKRTRIOGRVAHLTG 1200
 DB 1141 LSGANRLTLEDANIVQPLGLTILGKHLIYIDROQOMIERVEKTTGDKRTRIOGRVAHLTG 1200

QY 1201 IHAVEEVSLEEFSAHPCARDNGGCSHICIAKDGTPRCSCPVHLVLLQNLATCCGEPPTCS 1260
 DB 1201 IHAVEEVSLEEFSAHPCARDNGGCSHICIAKDGTPRCSCPVHLVLLQNLATCCGEPPTCS 1260

QY 1261 PDQFACATGEIDCIPGAWRCDGPFPECDQSDDEGCPVCSAAQFFPCARGCQVLDLRLCDGE 1320
 DB 1261 PDQFACATGEIDCIPGAWRCDGPFPECDQSDDEGCPVCSAAQFFPCARGCQVLDLRLCDGE 1320

QY 1321 ADCQDRSEADCAICLPNQPRCAGQCCLIKQCCDSFPDCIDGSDMLMCEBITKPPSDDS 1380
 DB 1321 ADCQDRSEADCAICLPNQPRCAGQCCLIKQCCDSFPDCIDGSDMLMCEBITKPPSDDS 1380

QY 1381 PAHSSAIGPVIGIILSLFVMGGVYFCQVVQRYAGANGFPHEYVSGTPHVPINFTAP 1440
 DB 1381 PAHSSAIGPVIGIILSLFVMGGVYFCQVVQRYAGANGFPHEYVSGTPHVPINFTAP 1440

QY 1441 GGSQHGPTGTIACCKSMSSVSLMGGRGVPLYDRNHVTGASSSSSSTKATLYPPIINP 1500
 DB 1441 GGSQHGPTGTIACCKSMSSVSLMGGRGVPLYDRNHVTGASSSSSSTKATLYPPIINP 1500

QY 1501 PPSPATPSLYNMDFYSSNIPATARYRPIYIRGMAPPPTPCSTDVCDSDYSASRWKAS 1560
 DB 1501 PPSPATPSLYNMDFYSSNIPATARYRPIYIRGMAPPPTPCSTDVCDSDYSASRWKAS 1560

QY 1561 KYILDLNSDSDPYPPPTPHSQVLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
 DB 1561 KYILDLNSDSDPYPPPTPHSQVLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 7

ADI27181

ID ADI27181 standard; protein; 1615 AA.

AC ADI27181;

DT 22-APR-2004 (first entry)

XX Human LRP binding family protein #12.

DE tooth development; Wnt pathway activation; bone deposition;

KW ocular development; cysteine knot; LRP binding family; human.

XX Homo sapiens.

XX

XX AC ADQ20524;
 XX DT 26-AUG-2004 (first entry)
 XX DE Human soft tissue sarcoma-upregulated protein - SEQ ID 3344.
 XX KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
 XX OS Homo sapiens.
 XX PN WO2004048938-A2.
 XX PD 10-JUN-2004.
 XX PF 26-NOV-2003; 2003WO-US038193.
 XX PR 26-NOV-2002; 2002US-0429739P.
 XX PA (PROT-) PROTEIN DESIGN LABS INC.
 XX PI Aziz N, Ginsburg WM, Zlotnik A;
 XX DR WPI; 2004-441208/41.
 XX PT Early detection of soft tissue sarcoma comprises determining expression
 PT of a gene in a first soft tissue sample and a normal soft tissue sample
 PT and comparing the gene expression, also useful in treating soft tissue
 PT sarcoma.
 XX PS Example 2; SEQ ID NO 3344; 210pp; English.
 XX CC The invention relates to a novel method for detecting soft tissue sarcoma
 CC which comprises obtaining a first soft tissue sample from an individual
 CC and a normal soft tissue sample from the same or different individual,
 CC determining the expression of a gene in both samples and comparing the
 CC expression of the gene in both soft tissue samples, where a higher level
 CC of protein expression in the first soft tissue sample indicates the
 CC presence of soft tissue sarcoma. The method of the invention has
 CC cytotatic applications and may be useful for detecting soft tissue
 CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
 CC acid sequences may be useful in diagnostic and screening applications.
 CC The current sequence is that of a human soft tissue sarcoma-upregulated
 CC protein of the invention. The current sequence is not shown within the
 CC specification per se but was submitted in CD format by the inventor.
 XX SQ Sequence 1615 AA;
 Query Match 100.0%; Score 8736; DB 8; Length 1615;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MEAAPGPPWPLLLLLLLALCGCPAPAAAPLFFANRRDRLVLDAGGVKLESTIVVS 60
 DB 1 MEAAPGPPWPLLLLLLLALCGCPAPAAAPLFFANRRDRLVLDAGGVKLESTIVVS 60
 QY 61 GLEDAAAVDFQSKGAVTWDVSEAIKQTYLNOTGAAVQNVVLSGLVSPDGLACDWVGK 120
 DB 61 GLEDAAAVDFQSKGAVTWDVSEAIKQTYLNOTGAAVQNVVLSGLVSPDGLACDWVGK 120
 QY 121 KLYWTDSETNRIEVANLNGTSRKVLFWODLQPPRAIALDPAHYMYWTDWGETPRIERAG 180
 DB 121 KLYWTDSETNRIEVANLNGTSRKVLFWODLQPPRAIALDPAHYMYWTDWGETPRIERAG 180
 QY 181 MDGSTRKIIVDSIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPRQKVVEGSLTTP 240
 DB 181 MDGSTRKIIVDSIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPRQKVVEGSLTTP 240
 QY 241 FALTLSGDTLYWTDWQTRSIIACNKRGTGGRKEILSALYSPMDIQVLSQERQPPFHTRCE 300
 DB 241 FALTLSGDTLYWTDWQTRSIIACNKRGTGGRKEILSALYSPMDIQVLSQERQPPFHTRCE 300
 QY 301 EDNGGCSHLCLLSPSEPFYTCACPTGVQLQNGRTCKAGAEVLLARRTDLRRISLDT 360

DB 301 EDNGGCSHLCLLSPSEPFYTCACPTGVQLQNGRTCKAGAEVLLARRTDLRRISLDT 360
 QY 361 DFTDIVLQVDDIRHAIADIDYDPLEGYVYVWTDDEVAIRRAYLDGSGAQTIVNTEINPDG 420
 DB 361 DFTDIVLQVDDIRHAIADIDYDPLEGYVYVWTDDEVAIRRAYLDGSGAQTIVNTEINPDG 420
 QY 421 IADVAVARNLYWTDGTDRIEVTRNLNGTSRKILSVSEDLDEPRALALHPVWGLMYWTDWGE 480
 DB 421 IADVAVARNLYWTDGTDRIEVTRNLNGTSRKILSVSEDLDEPRALALHPVWGLMYWTDWGE 480
 QY 481 NPKIECANLDGQERRVLVNASLGNPGLALDQSGKLYWGDATKTEVINVDGTKRRTL 540
 DB 481 NPKIECANLDGQERRVLVNASLGNPGLALDQSGKLYWGDATKTEVINVDGTKRRTL 540
 QY 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIEVHVHVKASRDVLIIDQLPDLMLGLKAVNAVAV 600
 DB 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIEVHVHVKASRDVLIIDQLPDLMLGLKAVNAVAV 600
 QY 601 GTNFCADRNGGCSHLCHFTPHATRCGCPIGLELLSDMKTCIIVPEAFIVFTSRAAIHRISL 660
 DB 601 GTNFCADRNGGCSHLCHFTPHATRCGCPIGLELLSDMKTCIIVPEAFIVFTSRAAIHRISL 660
 QY 661 ETNNNDVAIPLTGVKEASALDFDVSNHHIYWTDVSLKTIISRAFMNGSSVEHVHVEGLDYP 720
 DB 661 ETNNNDVAIPLTGVKEASALDFDVSNHHIYWTDVSLKTIISRAFMNGSSVEHVHVEGLDYP 720
 QY 721 EGMAVDMWGNKLYWADTGTNRIEVARLDGQPRQLVVRDLNPRSLALDPTKGYIYWTEW 780
 DB 721 EGMAVDMWGNKLYWADTGTNRIEVARLDGQPRQLVVRDLNPRSLALDPTKGYIYWTEW 780
 QY 781 GKKPRIIVRAFMDGTCNMTLVKGRANDLTIDYADQRLYWTDLDTNMIESNMIGQERVV 840
 DB 781 GKKPRIIVRAFMDGTCNMTLVKGRANDLTIDYADQRLYWTDLDTNMIESNMIGQERVV 840
 QY 841 IADLPPFPGLTQYSDIYWTDMNLHSIERADKTSGRNRTLIOGHLDFVMDILVFHSSRQ 900
 DB 841 IADLPPFPGLTQYSDIYWTDMNLHSIERADKTSGRNRTLIOGHLDFVMDILVFHSSRQ 900
 QY 901 DGLNDCMHNNQCQOLCLAIIPGHRCCASHYTLDPSSRNCSPPTTLLFFSQSAISAMI 960
 DB 901 DGLNDCMHNNQCQOLCLAIIPGHRCCASHYTLDPSSRNCSPPTTLLFFSQSAISAMI 960
 QY 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFYWVDGRQNIKRAKDDGTQPPVLSLSOGQ 1020
 DB 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFYWVDGRQNIKRAKDDGTQPPVLSLSOGQ 1020
 QY 1021 NPDRQPHDLSDIYSRITLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
 DB 1021 NPDRQPHDLSDIYSRITLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
 QY 1081 LYFTNMQDRAAKIBRAALDGTREVLFTTGLIRPALVVDNTLGLFWVDADLKRIESCD 1140
 DB 1081 LYFTNMQDRAAKIBRAALDGTREVLFTTGLIRPALVVDNTLGLFWVDADLKRIESCD 1140
 QY 1141 LSGANRLTLEDANIVQPLGTLILGKHYWTDROQMIERVEKTTGDKRTRIOGRVAHLTG 1200
 DB 1141 LSGANRLTLEDANIVQPLGTLILGKHYWTDROQMIERVEKTTGDKRTRIOGRVAHLTG 1200
 QY 1201 IHAVEEVSLEFSAPCAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQNLLTCGEPTPCS 1260
 DB 1201 IHAVEEVSLEFSAPCAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQNLLTCGEPTPCS 1260
 QY 1261 PDQFACATGEIDCIPGAWRCDGPFECDDQSDDEGCPVCSAAQPPCARGQCQVDLRLRCDGE 1320
 DB 1261 PDQFACATGEIDCIPGAWRCDGPFECDDQSDDEGCPVCSAAQPPCARGQCQVDLRLRCDGE 1320
 QY 1321 ADCQDRSDEADCAICLPNPFRCASGCVLTKQCDSPFDCIDGSDMLCMCITPPSPDS 1380
 DB 1321 ADCQDRSDEADCAICLPNPFRCASGCVLTKQCDSPFDCIDGSDMLCMCITPPSPDS 1380
 QY 1381 PAHSASLGPVIGITLSLFFVMGVVFCQRVVQRYAGANGPPFHYVSGTPIPLFIATP 1440

Db 1381 PAHSAIGPVIGIILSLFVGVYFVQVQVYAGANGPFPHEVSGTDPHVLNFIAP 1440
QY 1441 GSGHGGPTGTACGKSMSSVSLMGCGVPLXDRNHVTGASSSSSTKATLYPPIINP 1500
Db 1441 GSGHGGPTGTACGKSMSSVSLMGCGVPLXDRNHVTGASSSSSTKATLYPPIINP 1500
QY 1501 PPSPATPSLYNMDMFYSSNIPATAPYRPIIIRGMAPPPTTFCSTDVCDSDYSASRWKAS 1560
Db 1501 PPSPATPSLYNMDMFYSSNIPATAPYRPIIIRGMAPPPTTFCSTDVCDSDYSASRWKAS 1560
QY 1561 KYILDLSNDSPPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
Db 1561 KYILDLSNDSPPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 9

ADRI17560
ID ADRI17560 standard; protein; 1615 AA.

AC ADRI17560;
XX
DT 04-NOV-2004 (first entry)
XX
DE Human high bone mass gene, wild type allele Zmax1, protein #2.
XX
KW Human; high bone mass; Zmax1; HBM; osteoporosis; chromosome 11q13.3;
KW osteopathic; LDL receptor; bone development; metabolic bone disease; SNP;
KW single nucleotide polymorphism.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 171
FT /note= "May be Val as the result of a single nucleotide
polymorphism"

XX US6780609-B1.

XX 24-AUG-2004.

XX 05-APR-2000; 2000US-00543771.

XX 13-JAN-1998; 98US-0071449P.

XX 23-OCT-1998; 98US-0105511P.

XX 13-JAN-1999; 99US-00229319.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Carulli JP, Little RD, Recker RR, Johnson ML;
XX WPI; 2004-623529/60.

XX N-PSDB; ADRI16919.

XX New high bone mass gene of chromosome 11q13.3, encoding protein useful
PT for treating, diagnosing, preventing, or screening for normal and
PT abnormal conditions of bone, including metabolic bone diseases, e.g.
PT osteoporosis.

XX Disclosure; Fig 6; 284pp; English.

XX The invention relates to an isolated amino acid protein sequence selected
CC from an amino acid sequence appearing as ADRI16922 or an amino acid
CC sequence comprising or consisting of the extracellular domain of
CC ADRI16922 (amino acids 23-1385). ADRI16922 is encoded by the HBM (high bone
CC mass) allele of the human Zmax1 gene and has sequence similarity to LDL
CC receptors. Also disclosed are nucleic acids, proteins, cloning vectors,
CC expression vectors, transformed hosts, methods of developing
CC pharmaceutical compositions, methods of identifying molecules involved in
CC bone development, and methods of diagnosing and treating diseases
CC involved in bone development. Specifically disclosed is the Zmax1 gene
CC and the high bone mass (HBM) allele on chromosome 11q13.3 encoding
CC ADRI16922. The protein is useful for treating, diagnosing, preventing, or
CC screening for normal and abnormal conditions of bone, including metabolic

CC bone diseases, e.g. osteoporosis. The present sequence is the protein
CC encoded by the wild type human Zmax1 DNA. NOTE: Two versions of this
CC protein are present in the specification, ADRI16921 (sequence listing) and
CC ADRI17560 (shown as encoded by ADRI16919 in figure 6 and the sequence
CC listing). It is not clear which is the true protein since ADRI16921
CC contains translation exceptions and ADRI17560 doesn't.

XX Sequence 1615 AA;

Query Match 100.0%; Score 8736; DB 8; Length 1615;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEAAPPPPPPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRLLVDAGGVKLESTIVVS 60
Db 1 MEAAPPPPPPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRLLVDAGGVKLESTIVVS 60
QY 61 GLEDAAAVDFQFSKGVYVTDVSEBAIKQTYLNQTAGAQQVNVISGLVSPGLACDWGK 120
Db 61 GLEDAAAVDFQFSKGVYVTDVSEBAIKQTYLNQTAGAQQVNVISGLVSPGLACDWGK 120
QY 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLDPRAIALDPAHGYMYTWDGETPRIERAG 180
Db 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLDPRAIALDPAHGYMYTWDGETPRIERAG 180
QY 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPQKVVESGLTTP 240
Db 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPQKVVESGLTTP 240
QY 241 FALTSGDTLYTWDQTRSIIHACNKRITGGRKEILSALYSPMDIQVLSQERPPFHTRCE 300
Db 241 FALTSGDTLYTWDQTRSIIHACNKRITGGRKEILSALYSPMDIQVLSQERPPFHTRCE 300
QY 301 EDNGGCHLCLLSSEPPFYTCACPTGVQLQNGRTCKAGAEVLLARTRLSLDTTP 360
Db 301 EDNGGCHLCLLSSEPPFYTCACPTGVQLQNGRTCKAGAEVLLARTRLSLDTTP 360
QY 361 DFTDIVLQVDDIRHAIADYDPLGVYVTDDEVRARRAYLDGSGAOTLVNTEINDPDG 420
Db 361 DFTDIVLQVDDIRHAIADYDPLGVYVTDDEVRARRAYLDGSGAOTLVNTEINDPDG 420
QY 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVGLMYTWDGE 480
Db 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVGLMYTWDGE 480
QY 481 NPKIECANLDGQERRVLVNASLGNPGLADLQSGKLYWGDAKTDKIEVINVDGTKRRTL 540
Db 481 NPKIECANLDGQERRVLVNASLGNPGLADLQSGKLYWGDAKTDKIEVINVDGTKRRTL 540
QY 541 LEDKLPPIFGFTLLGDFIYTDWQRRSIEVHKVKASRDVIIDQLPDLMLGLKAVNVAKV 600
Db 541 LEDKLPPIFGFTLLGDFIYTDWQRRSIEVHKVKASRDVIIDQLPDLMLGLKAVNVAKV 600
QY 601 GTNPACDRNGGCSHLCFFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRISL 660
Db 601 GTNPACDRNGGCSHLCFFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRISL 660
QY 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWDVSLKTIISAPFMNGSSVEHVVEFGLDYP 720
Db 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWDVSLKTIISAPFMNGSSVEHVVEFGLDYP 720
QY 721 EGMAVDMMGNKLYWADGTNRIEVARLDGQPVLRDLNPRSLALDPTKGYIYTWTEW 780
Db 721 EGMAVDMMGNKLYWADGTNRIEVARLDGQPVLRDLNPRSLALDPTKGYIYTWTEW 780
QY 781 GKGPRIVRAFMDGTNCMTLVDKVGRANDLTIDYADQRLYWTDLDTNMISSNMLQERVV 840
Db 781 GKGPRIVRAFMDGTNCMTLVDKVGRANDLTIDYADQRLYWTDLDTNMISSNMLQERVV 840
QY 841 IADDLPHPPFGLTQVSDIYVTDWNLHLSIERADKTSGRNRTLIOGHLDPFMDILVFHSSRQ 900
Db 841 IADDLPHPPFGLTQVSDIYVTDWNLHLSIERADKTSGRNRTLIOGHLDPFMDILVFHSSRQ 900

QY 181 MDGSTRKIIIVDSIYWPNGLTIDLEEQKLYWADAKLSPIHRANLDGSPROKQVVGSLTHP 240
DB 181 MDGSTRKIIIVDSIYWPNGLTIDLEEQKLYWADAKLSPIHRANLDGSPROKQVVGSLTHP 240
QY 241 FALTLSGDTLWTDWQTRSIIHACNKRGTGGRKEILSALYSMDIQVLSQERQPPFHTRCE 300
DB 241 FALTLSGDTLWTDWQTRSIIHACNKRGTGGRKEILSALYSMDIQVLSQERQPPFHTRCE 300
QY 301 EDNGGCSHLCLLSPSEPYTCACPTGVOLQNGRTCKAGAEVLLAARTRTLRLSLDTP 360
DB 301 EDNGGCSHLCLLSPSEPYTCACPTGVOLQNGRTCKAGAEVLLAARTRTLRLSLDTP 360
QY 361 DFTDIVLQVDDIRHAIAIDYDPLEGYVYVTDDEVRARIRAYLDGSGAOTLVNTEINDPDG 420
DB 361 DFTDIVLQVDDIRHAIAIDYDPLEGYVYVTDDEVRARIRAYLDGSGAOTLVNTEINDPDG 420
QY 421 IAVDWARNLWTDGTDRIRIEVTRNLGTSRKILVSEDLDEPRALHHPVWGLMTWTDWGE 480
DB 421 IAVDWARNLWTDGTDRIRIEVTRNLGTSRKILVSEDLDEPRALHHPVWGLMTWTDWGE 480
QY 481 NPKIECANLDGQERRVLVNASLGNPGLALDLQEGKLYWDAKTDKIEVINVDGTKRRTL 540
DB 481 NPKIECANLDGQERRVLVNASLGNPGLALDLQEGKLYWDAKTDKIEVINVDGTKRRTL 540
QY 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIRVHKVKASRDVIIDQLPDLMLGLKANVAKV 600
DB 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIRVHKVKASRDVIIDQLPDLMLGLKANVAKV 600
QY 601 GTNPCADRNGGCSHLCPFTPHATRCGCPIGLELLSDMKTCIVPRAFLVFTSRAAIHRISL 660
DB 601 GTNPCADRNGGCSHLCPFTPHATRCGCPIGLELLSDMKTCIVPRAFLVFTSRAAIHRISL 660
QY 661 ETNNNDVAIPLTGVKEASALPDFVSNHHIYVTDVSLKTSIRAFWNGSSVEHVFGLDYP 720
DB 661 ETNNNDVAIPLTGVKEASALPDFVSNHHIYVTDVSLKTSIRAFWNGSSVEHVFGLDYP 720
QY 721 EGMADVMMGNKLYWADTGTNRIEVARLDGQFQVLVWRDLNPRSLALDPTKGYIYWTW 780
DB 721 EGMADVMMGNKLYWADTGTNRIEVARLDGQFQVLVWRDLNPRSLALDPTKGYIYWTW 780
QY 781 GKGPRIVRAFMVMDGNCMTLVKVGGRANDLTIDYADQRLYVTDLTNTNMISSNMLGQERV 840
DB 781 GKGPRIVRAFMVMDGNCMTLVKVGGRANDLTIDYADQRLYVTDLTNTNMISSNMLGQERV 840
QY 841 IADLPHFPFGITQYSDIYVTDNHLNLSIERADKTSGRNRTLIQGHLPVMDILVFHSSRQ 900
DB 841 IADLPHFPFGITQYSDIYVTDNHLNLSIERADKTSGRNRTLIQGHLPVMDILVFHSSRQ 900
QY 901 DGLNDCMHNNGCQGLCIAIPGHRGCGCASHYTLDPSSRNCSPPTTFLFQKSAISRMI 960
DB 901 DGLNDCMHNNGCQGLCIAIPGHRGCGCASHYTLDPSSRNCSPPTTFLFQKSAISRMI 960
QY 961 PDDQHSPLIILPLHGLRNVAIDYDPLDKFYVWDGRQNIKRAKDDGTQPPVLTSLSGQG 1020
DB 961 PDDQHSPLIILPLHGLRNVAIDYDPLDKFYVWDGRQNIKRAKDDGTQPPVLTSLSGQG 1020
QY 1021 NPDRQPHDLSIDIYSRTLFTCEATNTINVRHLSGEAMGVVLGRDRKPRAIIVNAERGY 1080
DB 1021 NPDRQPHDLSIDIYSRTLFTCEATNTINVRHLSGEAMGVVLGRDRKPRAIIVNAERGY 1080
QY 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKFWVDADLKRIESC 1140
DB 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKFWVDADLKRIESC 1140
QY 1141 LSGANRLTLEDANIVQPLGLTILGKHLVWIDRQQMIERVKTTGDKRTRIQGRVAHLT 1200
DB 1141 LSGANRLTLEDANIVQPLGLTILGKHLVWIDRQQMIERVKTTGDKRTRIQGRVAHLT 1200
QY 1201 IHAVEVSLSEFSAHPCARDNGGCSHICAKDGTGTPRSCPVHLVLLQNLITCGEPPTCS 1260
DB 1201 IHAVEVSLSEFSAHPCARDNGGCSHICAKDGTGTPRSCPVHLVLLQNLITCGEPPTCS 1260
QY 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEGCPVCSAAQFPCARQCQVDLRLRCDGE 1320

DB 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEGCPVCSAAQFPCARQCQVDLRLRCDGE 1320
QY 1321 ADCQDRSEADCAICLPNQPRCAGQCQLIKQOCDSPDCIDGSDLMCBITKPPSDDS 1380
DB 1321 ADCQDRSEADCAICLPNQPRCAGQCQLIKQOCDSPDCIDGSDLMCBITKPPSDDS 1380
QY 1381 PAHSSAIGPVGIIILSLFVMGVYVFCORVVCORVYAGANGPPHBYVSGTHVPLNFIAP 1440
DB 1381 PAHSSAIGPVGIIILSLFVMGVYVFCORVVCORVYAGANGPPHBYVSGTHVPLNFIAP 1440
QY 1441 GGSQHGPTGTIACGKMMSSVSLMGGRGVGLYDRNHVVTGASSSSSSSTKATLYPPIILNP 1500
DB 1441 GGSQHGPTGTIACGKMMSSVSLMGGRGVGLYDRNHVVTGASSSSSSSTKATLYPPIILNP 1500
QY 1501 PPSPATDPSLYNMDMFYSSNIPATARPYPVYIIRGMARPTTFCSTDVCDSDYSASRWKAS 1560
DB 1501 PPSPATDPSLYNMDMFYSSNIPATARPYPVYIIRGMARPTTFCSTDVCDSDYSASRWKAS 1560
QY 1561 KYLLDLNLSDDSPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
DB 1561 KYLLDLNLSDDSPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 11

ADR73482
ID ADR73482 standard; protein; 1615 AA.
XX
AC ADR73482;
XX
DT 02-DEC-2004 (first entry)
XX
DE Human low density lipoprotein receptor-related protein 5, LRP5, protein.
XX
KW LRP5; apoptosis; inhibitor; tumour; cytostatic; cancer.
XX
OS Homo sapiens.
XX
PN WO2004076682-A2.
XX
PD 10-SEP-2004.
XX
PF 26-FEB-2004; 2004WO-US0006020.
XX
PR 26-FEB-2003; 2003US-0450886P.
XX
PA (SURR-) SURROMED INC.
XX
PI Axenovich SA, Stull R, Gelman M, Chui K, Ng D;
XX
DR WPI; 2004-653428/63.
DR N-FSDB; ADR73481.
XX
PT Identifying a compound for inducing apoptosis, useful for diagnosing and
PT treating cancer, comprises identifying an inhibitor of a target.
XX
PS Claim 1; SEQ ID NO 50; 255pp; English.
XX
CC The invention relates to a novel method for identifying a compound for
CC inducing apoptosis comprising identifying an inhibitor of a target.
CC Examples of an inhibitor of a target include: angio-associated, migratory
CC cell protein (AAMP), disintegrin and metalloproteinase domain 8 (ADAM8),
CC disintegrin-like and metalloprotease (reporlysin type) with
CC thrombospondin type 1 motif, 17 (ADAMTS17), adenylate cyclase 3 (ADCY3),
CC adrenergic beta receptor kinase 1 (ADRBK1), bladder cancer associated
CC protein (BLCAP), chromosome 22 open reading frame 5 (C22orf5), CD81
CC antigen (target of antiproliferative antibody 1 (CD81), CD9 antigen (p24
CC (CD9), claudin 4 (CLDN4), chloride intracellular channel 1 (CLIC1),
CC collagen, type VI, alpha 2 (COL6A2), CTL2, endothelin converting enzyme 1
CC (ECE1), ephrinB1 (EPHB1), flotillin 2 (FLOT2), intercellular adhesion
CC molecule 3 (ICAM3), iduronate 2-sulfatase (Hunter syndrome) (IDS), jagged
CC 2 (JAG2), junctional adhesion molecule 1 (JAM1), lectin, galactoside-
CC binding soluble 3 binding protein (LGALS3BP), similar to possible G-

protein receptor (LOC146330), CGI-78 protein (LOC51107), lipoprotein lipase (LPL), low density lipoprotein receptor-related protein 5 (LRP5), Lutheran blood group (Auburger b antigen included) (LU), membrane component, chromosomell, surface marker 1 (MLIS1), serum constituent protein (MSE55), neuropathy target esterase (NTE), Homo sapiens CDNA FL31043 fig, clone HSYRA2000248 (PLEXIN A1) or Homo sapiens CDNA FLJ44113 fig, clone TESTI4046487, highly similar to Mus musculus plexin A1 (PLXNA1), protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF), interacting protein (liprin), alpha 3 (PPIA3), Homo sapiens peptide-histidine transporter 4 (PTR4), mRNA (PTR4), solute carrier family 16 (monocarboxylic acid transporters) member 3 (SLC16A3), solute carrier family 1 (neutral amino acid transporter) member 5 (SLC1A5), solute carrier family 39 (zinc transporter) member 3 (SLC39A1), serine protease inhibitor, Kunitz type 2 (SPINT2), stanniocalcin 2 (STC2), tumour necrosis receptor superfamily member 21 (TNFRSF21), tumour rejection antigen (gp96) 1 (TRAI), or transient receptor potential cation channel, subfamily M member 4 (TRPM4), respectively comprising any of the 40 sequences of 87-1615 amino acids (the even numbers selected from SEQ ID:2-80), shown in the specification. The invention further comprises a method for inducing apoptosis and a method for diagnosing a tumour. The novel compounds have cytostatic activity. The method is useful for identifying a compound for inducing apoptosis. The methods and compounds are useful for diagnosing and treating cancer, and in determining the prognosis for cancer in the patient or the susceptibility of the patient to a therapeutic treatment. This sequence represents one of the 40 protein targets of the invention.

CC protein receptor (LOC146330), CGI-78 protein (LOC51107), lipoprotein
 CC lipase (LPL), low density lipoprotein receptor-related protein 5 (LRP5),
 CC Lutheran blood group (Auburger b antigen included) (LU), membrane
 CC component, chromosomell, surface marker 1 (MLIS1), serum constituent
 CC protein (MSE55), neuropathy target esterase (NTE), Homo sapiens CDNA
 CC FL31043 fig, clone HSYRA2000248 (PLEXIN A1) or Homo sapiens CDNA
 CC fig, clone TESTI4046487, highly similar to Mus musculus plexin A1
 CC (PLXNA1), protein tyrosine phosphatase, receptor type, f polypeptide
 CC (PTPRF), interacting protein (liprin), alpha 3 (PPIA3), Homo sapiens
 CC peptide-histidine transporter 4 (PTR4), mRNA (PTR4), solute carrier
 CC family 16 (monocarboxylic acid transporters) member 3 (SLC16A3), solute
 CC carrier family 1 (neutral amino acid transporter) member 5 (SLC1A5),
 CC solute carrier family 39 (zinc transporter) member 3 (SLC39A1), serine
 CC protease inhibitor, Kunitz type 2 (SPINT2), stanniocalcin 2 (STC2),
 CC tumour necrosis receptor superfamily member 21 (TNFRSF21), tumour
 CC rejection antigen (gp96) 1 (TRAI), or transient receptor potential cation
 CC channel, subfamily M member 4 (TRPM4), respectively comprising any of the
 CC 40 sequences of 87-1615 amino acids (the even numbers selected from SEQ
 CC ID:2-80), shown in the specification. The invention further comprises a
 CC method for inducing apoptosis and a method for diagnosing a tumour. The
 CC novel compounds have cytostatic activity. The method is useful for
 CC identifying a compound for inducing apoptosis. The methods and compounds
 CC are useful for diagnosing and treating cancer, and in determining the
 CC prognosis for cancer in the patient or the susceptibility of the patient
 CC to a therapeutic treatment. This sequence represents one of the 40
 CC protein targets of the invention.

XX
 SQ Sequence 1615 AA;

Query Match 100.0%; Score 8736; DB 8; Length 1615;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEAAPGGPWPPLLLLLLLLLLALCGCPAPAPAAAPLPPANRRDRLVLDAGGVKLESTIVVS 60
 DB 1 MEAAPGGPWPPLLLLLLLLLLALCGCPAPAPAAAPLPPANRRDRLVLDAGGVKLESTIVVS 60
 QY 61 GLEDAADVDFQSKGAVYWTDSBEAIKQTYLNTGAAVQNVVISGLVSPDGLACDWGK 120
 DB 61 GLEDAADVDFQSKGAVYWTDSBEAIKQTYLNTGAAVQNVVISGLVSPDGLACDWGK 120
 QY 121 KLYWTDSTNRIEVANLNGTSRKVLFWQDLQDPAIALDPAHGYMYTWDGSETPRIERAG 180
 DB 121 KLYWTDSTNRIEVANLNGTSRKVLFWQDLQDPAIALDPAHGYMYTWDGSETPRIERAG 180
 QY 181 MDGSTRKLIIVSDIYWPNGLIDLEEQKLYWADAKLSFIHRANLDGSPRQKVGESLTHP 240
 DB 181 MDGSTRKLIIVSDIYWPNGLIDLEEQKLYWADAKLSFIHRANLDGSPRQKVGESLTHP 240
 QY 241 FALTLSGDTLYWTDQWTRSIHACNKRITGKKEILSALYSPMDIQVLSQERQPPFHTRCE 300
 DB 241 FALTLSGDTLYWTDQWTRSIHACNKRITGKKEILSALYSPMDIQVLSQERQPPFHTRCE 300
 QY 301 EDNGCSCHELLSPSEPPYTCACPTGVQLQNGRTCKAGAEVLLAARRTDLRLISLTP 360
 DB 301 EDNGCSCHELLSPSEPPYTCACPTGVQLQNGRTCKAGAEVLLAARRTDLRLISLTP 360
 QY 361 DFTDVLQVDDIRHAIAIDYDPLEGYVYWTDEVRIRRAYLDGSGAQTIVNTEINDPDG 420
 DB 361 DFTDVLQVDDIRHAIAIDYDPLEGYVYWTDEVRIRRAYLDGSGAQTIVNTEINDPDG 420
 QY 421 IAVDWARNLYWTDGTDRIEVTRLNGTSRKILYSEDLDEPRALALHPVGMGLMYTWDGGE 480
 DB 421 IAVDWARNLYWTDGTDRIEVTRLNGTSRKILYSEDLDEPRALALHPVGMGLMYTWDGGE 480
 QY 481 NPKIECANLQGOERVLVNASIGWPNGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL 540
 DB 481 NPKIECANLQGOERVLVNASIGWPNGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL 540
 QY 541 LEDKLPHIFGFTLLGDFLYWTDQWRRSIERVHKVYKASRDVLIIDQLPDLMLGLKAVNAVAVV 600
 DB 541 LEDKLPHIFGFTLLGDFLYWTDQWRRSIERVHKVYKASRDVLIIDQLPDLMLGLKAVNAVAVV 600

RESULT 12
 AAE21741

ID AAE21741 standard; protein; 1615 AA.

XX

QY 601 GTNPCADRNGCGSHLCFFTPHATRCGCPIGLELLSDMKTCTVPEAFIVFTSAAIHRISL 660
 DB 601 GTNPCADRNGCGSHLCFFTPHATRCGCPIGLELLSDMKTCTVPEAFIVFTSAAIHRISL 660
 QY 661 ETNNNDVAIPUTGVKEASALDFVSNHHIYWTDSVLSKTIISRAFMNGSSVEHVVEGLDYP 720
 DB 661 ETNNNDVAIPUTGVKEASALDFVSNHHIYWTDSVLSKTIISRAFMNGSSVEHVVEGLDYP 720
 QY 721 EGMADVMMGKNLYWADTGTNRIEVARLDGQFQVLRWDRDLNPRSLALDPTKGYIYWTEW 780
 DB 721 EGMADVMMGKNLYWADTGTNRIEVARLDGQFQVLRWDRDLNPRSLALDPTKGYIYWTEW 780
 QY 781 GKGPRIVAFMDGNCMTLVLDKVRANDLTIDYADQRLYWTDLNMTIENSSNMLGOERVV 840
 DB 781 GKGPRIVAFMDGNCMTLVLDKVRANDLTIDYADQRLYWTDLNMTIENSSNMLGOERVV 840
 QY 841 IADDLPHFPGLTQYSDIYWTDMNLHSIERADKTSGRNRTLIQGHLDPMVDLIVFHSRQ 900
 DB 841 IADDLPHFPGLTQYSDIYWTDMNLHSIERADKTSGRNRTLIQGHLDPMVDLIVFHSRQ 900
 QY 901 DGLNDGMHNGCCGQGLCIAIPGHRGCGASHYTLDPSRNGCSPPTTFLLFQSKAISRM 960
 DB 901 DGLNDGMHNGCCGQGLCIAIPGHRGCGASHYTLDPSRNGCSPPTTFLLFQSKAISRM 960
 QY 961 PDDQHSPLILPLHGLRNKALIDYDPLDKFIYWDGRQNIKRAKDDGTQPFVLTSLSGO 1020
 DB 961 PDDQHSPLILPLHGLRNKALIDYDPLDKFIYWDGRQNIKRAKDDGTQPFVLTSLSGO 1020
 QY 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
 DB 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
 QY 1081 LYFTNMQDRAAKIERAALDGTREVEVFTTGLIRPVALVVDNTLGLFWVDADLKRIESC 1140
 DB 1081 LYFTNMQDRAAKIERAALDGTREVEVFTTGLIRPVALVVDNTLGLFWVDADLKRIESC 1140
 QY 1141 LSGANRLTLEDAIVQPLGLTILGKLYWIDROQOMIERVEKTTGDKRTRIQGRVAHLTG 1200
 DB 1141 LSGANRLTLEDAIVQPLGLTILGKLYWIDROQOMIERVEKTTGDKRTRIQGRVAHLTG 1200
 QY 1201 IHAVEVSLEEFSAHPCARDNGCGSHICIAKGDGTPRCPCPVHLVLLQNLITCGEPPTCS 1260
 DB 1201 IHAVEVSLEEFSAHPCARDNGCGSHICIAKGDGTPRCPCPVHLVLLQNLITCGEPPTCS 1260
 QY 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEECPCVCSAAQFFCARGQCVDLRLCDE 1320
 DB 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEECPCVCSAAQFFCARGQCVDLRLCDE 1320
 QY 1321 ADCQDRDEADCDALCLPNQPCASGOCVLIKQCCDFPDCCI DGSDELMCEITKPPSDDS 1380
 DB 1321 ADCQDRDEADCDALCLPNQPCASGOCVLIKQCCDFPDCCI DGSDELMCEITKPPSDDS 1380
 QY 1381 PAHSSAIGPVIGITILSLFVMGGVYFCQVVCQRYAGANGFPFHEYVSGTGHVPLNFIA 1440
 DB 1381 PAHSSAIGPVIGITILSLFVMGGVYFCQVVCQRYAGANGFPFHEYVSGTGHVPLNFIA 1440
 QY 1441 GGSQHGPTGTACKSNMSSVSLMGGGVPLYDRNHVTGASSSSSSSTKATLYPPIINP 1500
 DB 1441 GGSQHGPTGTACKSNMSSVSLMGGGVPLYDRNHVTGASSSSSSSTKATLYPPIINP 1500
 QY 1501 PPSPATDPSLYNMDMFYSSNIPATAPRYPPIYIRGMAPPTTCTDVCDSYASRWKAS 1560
 DB 1501 PPSPATDPSLYNMDMFYSSNIPATAPRYPPIYIRGMAPPTTCTDVCDSYASRWKAS 1560
 QY 1561 KYIYLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
 DB 1561 KYIYLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

AAE21741;
 16-JUL-2002 (first entry)
 Human BSMR protein mutant, A1330L.
 Human; bone strength and mineralisation regulatory protein; BSMR;
 bone strength; mineralisation; ophthalmological; antidiabetic;
 bone density regulating transmembrane receptor; prosthetic device;
 surgical implant; diabetic retinopathy; hypertensive retinopathy;
 therapy; osteoporosis; prematurity; ocular vessel; eye disorder;
 osteopathic; mutant; mutin.
 Homo sapiens.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Key 1330
 FT Misc-difference 1330
 FT /note= "Wild-type Ala substituted with Leu"
 XX
 XX WO200216553-A2.
 XX
 XX 28-FEB-2002.
 PD
 XX
 XX 17-AUG-2001; 2001WO-US041788.
 XX
 XX 18-AUG-2000; 2000US-0226119P.
 PR
 XX 22-SEP-2000; 2000US-0234337P.
 PR
 XX 13-JUL-2001; 2001US-0304851P.
 XX
 XX (AVET) AVENTIS PHARMA SA.
 PA (HARD) HARVARD COLLEGE.
 PA (UYCA-) UNIV CASE WESTERN RESERVE.
 XX
 XX Warman ML, Gong Y, Olsen BR, Rawadi G, Roman-Roman S;
 PI WPI; 2002-329694/36.
 XX
 XX Polynucleotide encoding bone strength and mineralization regulatory
 PT protein useful for diagnosis or therapy of osteoporosis.
 PT
 XX Disclosure; Page; 124pp; English.
 PS
 XX
 XX The invention relates to bone strength and mineralisation regulatory
 CC protein (BSMR) and its corresponding nucleic acid sequence. BSMR DNA is
 CC useful for the diagnosis or therapy of osteoporosis and for regulating
 CC (increasing) bone strength and mineralisation in a human subject by
 CC activating a bone density regulating transmembrane receptor (BSMR
 CC protein). An expression vector comprising a promoter that is operably
 CC linked to BSMR DNA is useful for modulating bone density and for
 CC enhancing bone strength and mineralisation in a mammal cell. Composition
 CC comprising a BSMR effector is useful for treating osteoporosis and is
 CC useful particularly as a coating for prosthetic devices and surgical
 CC implants. BSMR is useful for screening lead pharmaceutical agents as BSMR
 CC effectors, which may be used to treat a range of eye disorders such as
 CC diabetic retinopathy, hypertensive retinopathy and retinopathy of
 CC prematurity, in which normal vascular growth and integrity of ocular
 CC vessels is disrupted. The present sequence is human BSMR protein mutant.
 CC Note: This sequence is not shown in the specification, however it is
 CC constructed based on the human BSMR protein shown in fig 3 of the
 CC specification (AAE21730)
 XX
 XX Sequence 1615 AA;
 SQ
 Query Match 99.9%; Score 8735; DB 5; Length 1615;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 1 MEAAPGPPWPLLLLLLLLLLALCGCPAPAAAPLFFANRRDRLVDAGGKLESTIVVS 60
 1 MEAAPGPPWPLLLLLLLLLLALCGCPAPAAAPLFFANRRDRLVDAGGKLESTIVVS 60
 61 GLEDAAAVDFQSKGAVVTDVSEAIKQTYLNTQGAQVNVISGLVSPDGLACDWGK 120

DB
 61 GLEDAAAVDFQSKGAVVTDVSEAIKQTYLNTQGAQVNVISGLVSPDGLACDWGK 120
 QY
 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDPRAIALDPAGHMYWTDWGETPRIERAG 180
 DB
 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDPRAIALDPAGHMYWTDWGETPRIERAG 180
 QY
 181 MDGSTRKLIIVDSDIYWPNGLTIDLEEOKLYWADAKLSFIHRANLDGSPQKQVVEGSLTHP 240
 DB
 181 MDGSTRKLIIVDSDIYWPNGLTIDLEEOKLYWADAKLSFIHRANLDGSPQKQVVEGSLTHP 240
 QY
 241 FALTSGDTLYWTDQTRSIIHACNKRITGGKKEILSALYSMDIQVLQSOERPPHTRCE 300
 DB
 241 FALTSGDTLYWTDQTRSIIHACNKRITGGKKEILSALYSMDIQVLQSOERPPHTRCE 300
 QY
 301 EDNGGCSHLCLLSPSEPFYTCACPTGVQDQNGRTCKAGAEVLLAARTDLRRLSLDTP 360
 DB
 301 EDNGGCSHLCLLSPSEPFYTCACPTGVQDQNGRTCKAGAEVLLAARTDLRRLSLDTP 360
 QY
 361 DFTDIVLQVDDIRHAIADYDPLEGYVYVTDDEVRARRAYLDGSGAQTLYNTEINDPDG 420
 DB
 361 DFTDIVLQVDDIRHAIADYDPLEGYVYVTDDEVRARRAYLDGSGAQTLYNTEINDPDG 420
 QY
 421 IAVDVARNLWYTDGTDRIEVTRNGTSRKILVSEDLDEPRALALHPVWGLMYTWDGE 480
 DB
 421 IAVDVARNLWYTDGTDRIEVTRNGTSRKILVSEDLDEPRALALHPVWGLMYTWDGE 480
 QY
 481 NPKTECANLQGERRLVNASLGNPGLALDQSGKLYWDAKTKDIEVINVDGTKRRTL 540
 DB
 481 NPKTECANLQGERRLVNASLGNPGLALDQSGKLYWDAKTKDIEVINVDGTKRRTL 540
 QY
 541 LEDKLPHIFGFTLLGDFIYTDWQRRSIRVHKVKSADVIIDOLPDLMLKAVNAVKV 600
 DB
 541 LEDKLPHIFGFTLLGDFIYTDWQRRSIRVHKVKSADVIIDOLPDLMLKAVNAVKV 600
 QY
 601 GTNFCADRNGCSHLCEFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRLS 660
 DB
 601 GTNFCADRNGCSHLCEFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRLS 660
 QY
 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWDVSLTKTISRPFNGSSVEHVVEGLDYP 720
 DB
 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWDVSLTKTISRPFNGSSVEHVVEGLDYP 720
 QY
 721 EGMADVMMGNKLYWADGTNRIEVARLDGQPROVLVWRDLNPRSLALDPTKGYIYWTW 780
 DB
 721 EGMADVMMGNKLYWADGTNRIEVARLDGQPROVLVWRDLNPRSLALDPTKGYIYWTW 780
 QY
 781 GGGPRIIVRAFMDGTNCMTLVDKVGRANDLTIDYADQRLYWTDLDTNMISSNMLQGERVV 840
 DB
 781 GGGPRIIVRAFMDGTNCMTLVDKVGRANDLTIDYADQRLYWTDLDTNMISSNMLQGERVV 840
 QY
 841 IADLPHPPFGLTQYSDIYIYTDWNLHSIERADKTSGRNRTLIQGHLPDWMILVPHSSRQ 900
 DB
 841 IADLPHPPFGLTQYSDIYIYTDWNLHSIERADKTSGRNRTLIQGHLPDWMILVPHSSRQ 900
 QY
 901 DGLNDCHNNNGCCGOLCLAIPEGHRCGCASHYTLDPSSNCSPPTTFLFSOKSAISRWI 960
 DB
 901 DGLNDCHNNNGCCGOLCLAIPEGHRCGCASHYTLDPSSNCSPPTTFLFSOKSAISRWI 960
 QY
 961 PDDQHSPLDILPLHGLNRVKAIDYDPLDKFTYVWDGQNIKRAKDDGTQPPFVLTSLSQGQ 1020
 DB
 961 PDDQHSPLDILPLHGLNRVKAIDYDPLDKFTYVWDGQNIKRAKDDGTQPPFVLTSLSQGQ 1020
 QY
 1021 NPDQPHDLSDIYSRITLFWTCEATNTINVHRLSGEAMGVVLRGRDRKPRAIIVNAERG 1080
 DB
 1021 NPDQPHDLSDIYSRITLFWTCEATNTINVHRLSGEAMGVVLRGRDRKPRAIIVNAERG 1080
 QY
 1081 LYFTNMQRRAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESCD 1140
 DB
 1081 LYFTNMQRRAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESCD 1140
 QY
 1141 LSGANRLTLEDANIVQPLGLTILGKHLIYIDRQQQMTERVEKTKDKRTRIQGRVAHLTG 1200

Db 1141 LSGANRLTEDANIVQPLGLTILGKHLWIDRQQQMIERVEKTTGDKRTRIQGRVAHLTG 1200
 QY 1201 IHAVEVLEBFAHPCARDNGGCSHICIAKGDGTPRCSCPVHLVLLQNLITCGEPPTCS 1260
 Db 1201 IHAVEVLEBFAHPCARDNGGCSHICIAKGDGTPRCSCPVHLVLLQNLITCGEPPTCS 1260
 QY 1261 PDQFACATGEIDICPGAWRCDFPCCDDQSBEGCPVCSAAQFPCCARGCQVDLRLRCDDGE 1320
 Db 1261 PDQFACATGEIDICPGAWRCDFPCCDDQSBEGCPVCSAAQFPCCARGCQVDLRLRCDDGE 1320
 QY 1321 ADCQDRSEADCAICLENQFRCASGQCVLIKQCCDSFPDCIDGSDLMCEITKPPSDDS 1380
 Db 1321 ADCQDRSEADCAICLENQFRCASGQCVLIKQCCDSFPDCIDGSDLMCEITKPPSDDS 1380
 QY 1381 PAHSSAIGPVGIIILSLFVMGVYFVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 1440
 Db 1381 PAHSSAIGPVGIIILSLFVMGVYFVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 1440
 QY 1441 GGSQHGPTGTACGKSMSSVSLMGGRGVPLVYDRNVHTVGASSSSSSSTKATLYPPILNP 1500
 Db 1441 GGSQHGPTGTACGKSMSSVSLMGGRGVPLVYDRNVHTVGASSSSSSSTKATLYPPILNP 1500
 QY 1501 PPSPATPSLYNMDMFYSSNIPATAPRYPIIRGMAPPPTTPTCSTDVCDSDYSASRWKAS 1560
 Db 1501 PPSPATPSLYNMDMFYSSNIPATAPRYPIIRGMAPPPTTPTCSTDVCDSDYSASRWKAS 1560
 QY 1561 KYILDNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
 Db 1561 KYILDNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 13

ADRI17561
 ID ADRI17561 standard; protein; 1615 AA.
 AC ADRI17561;
 XX
 DT 04-NOV-2004 (first entry)
 DE Human high bone mass gene, HBM allele, protein #2.
 KW Human; high bone mass; Zmax1; HBM; osteoporosis; chromosome 11q13.3;
 KW osteopathic; LDL receptor; bone development; metabolic bone disease; SNP;
 KW single nucleotide polymorphism.
 XX
 OS Homo sapiens.
 FH
 FT Key Location/Qualifiers
 FT Misc-difference 171
 FT /note= "May be Gly as the result of a single nucleotide
 FT polymorphism"
 XX
 FN US6780609-B1.
 PD 24-AUG-2004.
 XX
 PF 05-APR-2000; 2000US-00543771.
 PR 13-JAN-1998; 98US-0071449P.
 PR 23-OCT-1998; 98US-0105511P.
 PR 13-JAN-1999; 99US-00229319.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Carulli JP, Little RD, Recker RR, Johnson ML;
 DR WPI; 2004-623529/60.
 DR N-PSDB; ADRI16920.
 XX
 PT New high bone mass gene of chromosome 1.1Q13.3, encoding protein useful
 PT for treating, diagnosing, preventing, or screening for normal and
 PT abnormal conditions of bone, including metabolic bone diseases, e.g.
 PT osteoporosis.

XX PS Disclosure; Col 99-110; 284pp; English.
 CC The invention relates to an isolated amino acid protein sequence selected
 CC from an amino acid sequence appearing as ADRI16922 or an amino acid
 CC sequence comprising or consisting of the extracellular domain of
 CC ADRI16922(amino acids 23-1385). ADRI16922 is encoded by the HBM (high bone
 CC mass) allele of the human Zmax1 gene and has sequence similarity to LDL
 CC receptors. Also disclosed are nucleic acids, proteins, cloning vectors,
 CC expression vectors, transformed hosts, methods of developing
 CC pharmaceutical compositions, methods of identifying molecules involved in
 CC bone development, and methods of diagnosing and treating diseases
 CC involved in bone development. Specifically disclosed is the Zmax1 gene
 CC and the high bone mass (HBM) allele on chromosome 11q13.3 encoding
 CC ADRI16922. The protein is useful for treating, diagnosing, preventing, or
 CC screening for normal and abnormal conditions of bone, including metabolic
 CC bone diseases, e.g. osteoporosis. The present sequence is the protein
 CC encoded by the disease causing allele of human Zmax1. HBM. NOTE: Two
 CC versions of this protein are present in the specification, ADRI16922
 CC (sequence listing) and ADRI17561 (shown as encoded by ADRI16920 in the
 CC sequence listing). It is not clear which is the true protein since
 CC ADRI16922 contains translation exceptions and ADRI17561 doesn't.
 XX SQ Sequence 1615 AA;
 Query Match 99.9%; Score 8727; DB 8; Length 1615;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1613; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MEAAPPPPPPLLLLLLLLLLALCCGAPAPAAAPLPPANRRDVRVLDAGGVKLESTIVVS 60
 Db 1 MEAAPPPPPPLLLLLLLLLLALCCGAPAPAAAPLPPANRRDVRVLDAGGVKLESTIVVS 60
 QY 61 GLEDAAAVDFQFSKGVAVTVDVSEAIKQTVLNTGAAVQNVVLSGLVSPGLACDWGK 120
 Db 61 GLEDAAAVDFQFSKGVAVTVDVSEAIKQTVLNTGAAVQNVVLSGLVSPGLACDWGK 120
 QY 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLDPRAIALDPAGHYMTWTDWCTETRIERAG 180
 Db 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLDPRAIALDPAGHYMTWTDWCTETRIERAG 180
 QY 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSGFRQKVVEGSLTTP 240
 Db 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSGFRQKVVEGSLTTP 240
 QY 241 FALTLSGDTLYWTDWQTRS IHACNKRITGGKKEILSALYSPMDIQVLSQERQPPFHTTCE 300
 Db 241 FALTLSGDTLYWTDWQTRS IHACNKRITGGKKEILSALYSPMDIQVLSQERQPPFHTTCE 300
 QY 301 EDNGCSHLCLLSPSEFPYTCACPTGVQLQDNGRTCKAGAEVLLARTRDLRLSLDTP 360
 Db 301 EDNGCSHLCLLSPSEFPYTCACPTGVQLQDNGRTCKAGAEVLLARTRDLRLSLDTP 360
 QY 361 DFTDIVLQVDDIRHAIADIDVPLEGYVYVWTDDEVRAIRRAYLDGSGAQLVNTINDPDG 420
 Db 361 DFTDIVLQVDDIRHAIADIDVPLEGYVYVWTDDEVRAIRRAYLDGSGAQLVNTINDPDG 420
 QY 421 IAVDWARNLYWTDGTDRIEVTRNLGTSRKILVSDELDEPRALHPVGLMYTWDWGE 480
 Db 421 IAVDWARNLYWTDGTDRIEVTRNLGTSRKILVSDELDEPRALHPVGLMYTWDWGE 480
 QY 481 NPKIECANLDGQERRVLVNASLGNPLALDQSGKLYWDAKTDKIEVINVDGTKRRTL 540
 Db 481 NPKIECANLDGQERRVLVNASLGNPLALDQSGKLYWDAKTDKIEVINVDGTKRRTL 540
 QY 541 LEDKLPHIFGFTLLGDFIYWTWQRRSIEHVHVKASRDVLIIDQLPDLMLGLKANVAKV 600
 Db 541 LEDKLPHIFGFTLLGDFIYWTWQRRSIEHVHVKASRDVLIIDQLPDLMLGLKANVAKV 600
 QY 601 GTNFCADNRGGCSHLCPFTTHATECGCPIGLELLSDMKTCLVPEAFVFTSRAAHRISL 660
 Db 601 GTNFCADNRGGCSHLCPFTTHATECGCPIGLELLSDMKTCLVPEAFVFTSRAAHRISL 660

Query Match		99.9%; Score 8727; DB 8; Length 1615;																																
Best Local Similarity		99.9%; Pred. No. 0;																																
Matches 1613; Conservative		0; Mismatches 2; Indels 0; Gaps 0;																																
QY	1	MEAAPGGPWP	LLLLLLLLLL	CGCPA	PAAASPLLL	FANRRDVR	LVDAGGVKLESTIYVS 60																											
Db	1	MEAAPGGPWP	LLLLLLLLLL	CGCPA	PAAASPLLL	FANRRDVR	LVDAGGVKLESTIYVS 60																											
QY	61	GLEDAADV	QFQSKAVY	WTDVSE	AIKQTYL	NOTGA	VQNVVISGLVSPDGLACDWGK 120																											
Db	61	GLEDAADV	QFQSKAVY	WTDVSE	AIKQTYL	NOTGA	VQNVVISGLVSPDGLACDWGK 120																											
QY	121	KLYWTDSE	TNRLEV	ANLNGT	SRKVL	FWQDLO	PRALALDPAGHYMYWTDWGETPRIERAG 180																											
Db	121	KLYWTDSE	TNRLEV	ANLNGT	SRKVL	FWQDLO	PRALALDPAGHYMYWTDWGETPRIERAG 180																											
QY	181	MDGSTRKI	IIVSDI	YWPNGLT	IDLE	EQKLY	WADAKLSFIHRANLDGSRQKVVESGLTHP 240																											
Db	181	MDGSTRKI	IIVSDI	YWPNGLT	IDLE	EQKLY	WADAKLSFIHRANLDGSRQKVVESGLTHP 240																											
QY	241	FALTLSG	TLWTDQ	TRSIH	ACNKR	TGKKE	ILSALYSPMDIQVLQOERQPPFHTRCE 300																											
Db	241	FALTLSG	TLWTDQ	TRSIH	ACNKR	TGKKE	ILSALYSPMDIQVLQOERQPPFHTRCE 300																											
QY	301	EDNGGCS	HLCLLS	PSPEP	VTCA	PTGVOL	QDNRTCKAGAEVILLARRTDLRLISLDT 360																											
Db	301	EDNGGCS	HLCLLS	PSPEP	VTCA	PTGVOL	QDNRTCKAGAEVILLARRTDLRLISLDT 360																											
QY	361	DFTDVLQ	VQDDIR	HAIAID	YDP	LEGVY	WTDDEVRAIRRAYLDGSGAQTLVNTEINDPDG 420																											
Db	361	DFTDVLQ	VQDDIR	HAIAID	YDP	LEGVY	WTDDEVRAIRRAYLDGSGAQTLVNTEINDPDG 420																											
QY	421	IADVARN	LYWTDG	TRIEV	TRNGT	SRKLS	VEDLDEPRALHPVGMGLMYWTDWGE 480																											
Db	421	IADVARN	LYWTDG	TRIEV	TRNGT	SRKLS	VEDLDEPRALHPVGMGLMYWTDWGE 480																											
QY	481	NPKIECAN	LDQERR	VLYN	ASL	GWENGL	ALDLQEGKLYWDAKTDKIEVINVDGTKRRTL 540																											
Db	481	NPKIECAN	LDQERR	VLYN	ASL	GWENGL	ALDLQEGKLYWDAKTDKIEVINVDGTKRRTL 540																											
QY	541	LEDKLP	PHIFG	TLGDF	YIY	WTDQ	RRSIRERHVKVKA	SRDVIIDQLPDLMLKAVNVAKV 600																										
Db	541	LEDKLP	PHIFG	TLGDF	YIY	WTDQ	RRSIRERHVKVKA	SRDVIIDQLPDLMLKAVNVAKV 600																										
QY	601	GTNP	CADRNG	CGSHL	CFT	PHATRC	GCPIG	LELLSDMKTCIVPEAFVFTSRAIHRISL 660																										
Db	601	GTNP	CADRNG	CGSHL	CFT	PHATRC	GCPIG	LELLSDMKTCIVPEAFVFTSRAIHRISL 660																										
QY	661	ETNNND	VAIPL	TGVKE	ASAL	DFD	VSNNHI	YWTDVSLKTI	SRAFMNGSSVEHVFEGLDYP 720																									
Db	661	ETNNND	VAIPL	TGVKE	ASAL	DFD	VSNNHI	YWTDVSLKTI	SRAFMNGSSVEHVFEGLDYP 720																									
QY	721	EGMAV	DMWGN	KLYW	ADT	GTNR	IEVAR	LDGQ	FRQVLVWRDLNPRSLALDPTKGYIYWTEW 780																									
Db	721	EGMAV	DMWGN	KLYW	ADT	GTNR	IEVAR	LDGQ	FRQVLVWRDLNPRSLALDPTKGYIYWTEW 780																									
QY	781	GGKPI	VPAF	MDG	TNC	MTLV	DKVGR	ANDLT	DIYADQRLY	WTDLTDTNMTI	ESSNMLQ	QERRV 840																						
Db	781	GGKPI	VPAF	MDG	TNC	MTLV	DKVGR	ANDLT	DIYADQRLY	WTDLTDTNMTI	ESSNMLQ	QERRV 840																						
QY	841	IADDL	PHFP	GLT	QYSD	YIY	WTDN	LHSI	ERADK	TSGR	NRTL	IQHLD	FVMD	ILVPHSSRQ 900																				
Db	841	IADDL	PHFP	GLT	QYSD	YIY	WTDN	LHSI	ERADK	TSGR	NRTL	IQHLD	FVMD	ILVPHSSRQ 900																				
QY	901	DGLND	CMHNG	CGGOL	CLAI	PG	HR	CGC	ASHY	TLDP	SSRNC	SPPT	FTLL	FP	SQKSA	ISRM 960																		
Db	901	DGLND	CMHNG	CGGOL	CLAI	PG	HR	CGC	ASHY	TLDP	SSRNC	SPPT	FTLL	FP	SQKSA	ISRM 960																		
QY	961	PDDQHS	PD	LIL	PL	HGL	RNV	KALD	YD	PL	DKFY	WVD	GRNI	KRA	DKD	GT	Q	FP	V	L	S	LS	Q	Q	1020									
Db	961	PDDQHS	PD	LIL	PL	HGL	RNV	KALD	YD	PL	DKFY	WVD	GRNI	KRA	DKD	GT	Q	FP	V	L	S	LS	Q	Q	1020									
QY	1021	NPDRQ	PHDLS	DI	YS	RTL	FW	TCE	ATNT	IN	VH	LS	GE	AM	GV	V	LR	LD	GR	DR	K	P	R	A	I	V	V	N	A	E	R	G	Y	1080

RESULT 15

ABM85665

ID ABM85665 standard; protein; 1627 AA.

XX AC ABM85665;

XX DT 18-NOV-2004 (first entry)

XX DE Human protein sequence hCP1690976.

XX KW Cytostatic; carcinoma; lymphoma; cancer; human.

XX OS Homo sapiens.

XX PN WO2003073826-A2.

XX PD 12-SEP-2003.

XX PF 28-FEB-2003; 2003WO-US006235.

XX PR 01-MAR-2002; 2002US-00087192.

XX PA (SAGR-) SAGRES DISCOVERY.

XX PI Morris DW;

XX DR WPI; 2003-328604/31.

XX PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma

XX PT comprises a nucleotide sequence.

XX PS Claim 5; SEQ ID NO 1410; Opp; English.

XX XX

CC The present invention relates to novel DNA and protein sequences which
 CC are associated with carcinomas. The sequences are useful for: (i) for
 CC screening drug candidates; (ii) for screening of bioactive agent capable
 CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
 CC a bioactive agent capable of modulating the activity of CAP; (iv) for
 CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
 CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
 CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
 CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
 CC determining Carcinoma Associated (CA) gene copy number. In addition, the
 CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
 CC carcinoma including lymphoma. The present sequence is one such CAP. Note:
 CC This patent is an equivalent to basic patent US2002182586A1, for which no
 CC sequence data was published

XX Sequence 1627 AA;

Query Match 99.8%; Score 8724; DB 7; Length 1627;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1615; Conservative 0; Mismatches 0; Indels 12; Gaps 1;

Qy	1	MEAPPPGPPWPLLLLLLLALCGCPAPAAASPLLFANRRDVRVLVDAGGVKLESTIVVS	60
Db	1	MEAPPPGPPWPLLLLLLLALCGCPAPAAASPLLFANRRDVRVLVDAGGVKLESTIVVS	60
Qy	61	GLEDAAAVDFQFSKGAVTWTVSEBAIKQTYLNOTGAQVQNVVISGLVSPDGLACDWYVK	120
Db	61	GLEDAAAVDFQFSKGAVTWTVSEBAIKQTYLNOTGAQVQNVVISGLVSPDGLACDWYVK	120
Qy	121	KLYWTDSETNRIEVANLNGTSRKVLFWQDLOQRAIALDPAHGYMYTWDGTPRIERAG	180
Db	121	KLYWTDSETNRIEVANLNGTSRKVLFWQDLOQRAIALDPAHGYMYTWDGTPRIERAG	180
Qy	181	MDGSTRKLIIVDSIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPROKVVGSLTTP	240
Db	181	MDGSTRKLIIVDSIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPROKVVGSLTTP	240
Qy	241	FALTLSGDTLYWTDQWTRS IHACNKTGKGRKEILSALYSPMDIQVLSQERQPPFHTRCE	300
Db	241	FALTLSGDTLYWTDQWTRS IHACNKTGKGRKEILSALYSPMDIQVLSQERQPPFHTRCE	300
Qy	301	EDNGCCHLCLLSPEPPTACPTGVOLQNGRTCKAGAEVLLAARTRDLRISLDP	360
Db	301	EDNGCCHLCLLSPEPPTACPTGVOLQNGRTCKAGAEVLLAARTRDLRISLDP	360
Qy	361	DFTDVLQVDDIRHAIAIDYDPLEGVYVWTDDEVRAIRRAYLDGSGAOTLVNTEINDPDG	420
Db	361	DFTDVLQVDDIRHAIAIDYDPLEGVYVWTDDEVRAIRRAYLDGSGAOTLVNTEINDPDG	420
Qy	421	IADVVAARNLYWTDGTDRIEVTRNLNGTSRKILVSEDLDEPRAIALHPVMGLMYTWDGE	480
Db	421	IADVVAARNLYWTDGTDRIEVTRNLNGTSRKILVSEDLDEPRAIALHPVMGLMYTWDGE	480
Qy	481	NPKIECANLDOQRRVLVNASLGNWGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL	540
Db	481	NPKIECANLDOQRRVLVNASLGNWGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL	540
Qy	541	LEDKLPHTFGTLLGDFLYWTDQWRRSIEVHKVKAERDVIIIDQLPDLMLKAVNAVAVV	600
Db	541	LEDKLPHTFGTLLGDFLYWTDQWRRSIEVHKVKAERDVIIIDQLPDLMLKAVNAVAVV	600
Qy	601	GTNPCADRNGGCHLCFFTPHATRCGCPGLELLSDMTKCIVPEAFVFTSRAAIIHRLS	660
Db	601	GTNPCADRNGGCHLCFFTPHATRCGCPGLELLSDMTKCIVPEAFVFTSRAAIIHRLS	660
Qy	661	ETNNNDVAIPLTGVEASALDFVSNHHIYWTDVSLKTSIRAFMNGSSVEHVFEGLDYP	720
Db	661	ETNNNDVAIPLTGVEASALDFVSNHHIYWTDVSLKTSIRAFMNGSSVEHVFEGLDYP	720
Qy	721	EGMAVDWNGKLYWADTGTNRIEVARLDGQFQVILVWRDLNPRSLALDPTKGYIYWTEW	780
Db	721	EGMAVDWNGKLYWADTGTNRIEVARLDGQFQVILVWRDLNPRSLALDPTKGYIYWTEW	780

Qy	781	GGKPRIVRAFMGDMTCMTLVKVGRLANDLTIDYDQRLYWTDLTNTMISSNMLQERVV	840
Db	781	GGKPRIVRAFMGDMTCMTLVKVGRLANDLTIDYDQRLYWTDLTNTMISSNMLQERVV	840
Qy	841	IADDLPHFPGLTQYSDVIYWTDMNLHSIERADKTSGRNRTLIQGLDPMVILVPHSSRQ	900
Db	841	IADDLPHFPGLTQYSDVIYWTDMNLHSIERADKTSGRNRTLIQGLDPMVILVPHSSRQ	900
Qy	901	DGLNDCHNNGCQCQLCLAIIPGGRHRCGCASHYTLDPSSRNCSPPTTFLLFQKSAISMI	960
Db	901	DGLNDCHNNGCQCQLCLAIIPGGRHRCGCASHYTLDPSSRNCSPPTTFLLFQKSAISMI	960
Qy	961	PDDQHSPLILPLHGLRNVAKAIDYDPLDKFYIYWDGRQNIKRAKDDGT-----	1008
Db	961	PDDQHSPLILPLHGLRNVAKAIDYDPLDKFYIYWDGRQNIKRAKDDGTQAVALSLGFFVL	1020
Qy	1009	OPFVLTSLSQONPDQPHDLSIDIYSRTLTFTWTEATNTINVHRLSGEAMGVLRGDRDK	1068
Db	1021	OPFVLTSLSQONPDQPHDLSIDIYSRTLTFTWTEATNTINVHRLSGEAMGVLRGDRDK	1080
Qy	1069	PRAIVVNAERGYLEFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKFW	1128
Db	1081	PRAIVVNAERGYLEFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKFW	1140
Qy	1129	VDADLKRIESCDLSGANRLTLEDANI VQPLGLTILGKHLIYWDQQQMIERVEKTTGDKR	1188
Db	1141	VDADLKRIESCDLSGANRLTLEDANI VQPLGLTILGKHLIYWDQQQMIERVEKTTGDKR	1200
Qy	1189	TRIOGRAVHLTGIIHAEVSEVSELEFSNAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQ	1248
Db	1201	TRIOGRAVHLTGIIHAEVSEVSELEFSNAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQ	1260
Qy	1249	NLLTCGGPPTCSPPQFACATGEIDCI PGAWRCDFPECCDDQDEBEGCPVCSAAQPPCARG	1308
Db	1261	NLLTCGGPPTCSPPQFACATGEIDCI PGAWRCDFPECCDDQDEBEGCPVCSAAQPPCARG	1320
Qy	1309	QCVDLRLRCDEADCDQRSDEADCAICLPNQRCASQCQVLI KQCCDSFDCIDGSDCL	1368
Db	1321	QCVDLRLRCDEADCDQRSDEADCAICLPNQRCASQCQVLI KQCCDSFDCIDGSDCL	1380
Qy	1369	MCEITKPPSDSPAHSSAIGPVGIIILSLFVNGGVYFVCQVVCQRYAGANGPPPEHYVS	1428
Db	1381	MCEITKPPSDSPAHSSAIGPVGIIILSLFVNGGVYFVCQVVCQRYAGANGPPPEHYVS	1440
Qy	1429	GTPHVPLNFIAPGSGOHPPTGIIACGKSMSSVSLMGGRGVPLYDRNHVTGASSSSSS	1488
Db	1441	GTPHVPLNFIAPGSGOHPPTGIIACGKSMSSVSLMGGRGVPLYDRNHVTGASSSSSS	1500
Qy	1489	TKATLYPPIILNPPSPATDPSLYNMDFYSSNIPATAPRYPIYIRGMAPPPTPCSTDVC	1548
Db	1501	TKATLYPPIILNPPSPATDPSLYNMDFYSSNIPATAPRYPIYIRGMAPPPTPCSTDVC	1560
Qy	1549	DSYASASRWKASKYLYDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLFPPPP	1608
Db	1561	DSYASASRWKASKYLYDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLFPPPP	1620
Qy	1609	SPCTDSS 1615	
Db	1621	SPCTDSS 1627	

Search completed: February 17, 2005, 01:31:02
 Job time : 210 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2005, 01:16:38 ; Search time 224 Seconds
(without alignments)
3691.999 Million cell updates/sec

Title: US-09-931-375A-2

Perfect score: 8740

Sequence: 1 MEAAPGPPWPLLLLLLLLL.....TERSYFHLPPPPSPCTDSS 1615

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8740	100.0	1615	2 QUES7	Ques7 homo sapien
2	8736	100.0	1615	1 LRP5 HUMAN	Q75197 homo sapien
3	8259.5	94.5	1614	1 LRP5_MOUSE	Q91v00 mus musculus
4	7132.5	81.6	1605	2 Q8AYF1	Q8AYF1 xenopus lae
5	6153.5	70.4	1613	1 LRP6_HUMAN	Q75581 homo sapien
6	6097.5	69.8	1613	1 LRP6_MOUSE	Q88572 mus musculus
7	6084	69.6	1613	2 Q8AYF0	Q8AYF0 xenopus lae
8	3527.5	40.4	1678	2 Q9V609	Q95v09 drosophila
9	3526.5	40.3	1678	2 Q9V600	Q9v600 drosophila
10	3521.5	40.3	1678	2 Q9NHE9	Q9nhes drosophila
11	3262.5	37.3	1698	2 Q7PV65	Q7pv65 anopheles g
12	2788	31.9	1246	2 Q6AWJ8	Q6awj8 drosophila
13	2687	30.7	1905	1 LRP4_RAT	Q9qypl rattus norv
14	2687	30.7	1905	2 Q76LJ2	Q76lj2 rattus norv
15	2686	30.7	1950	1 LRP4_HUMAN	Q75096 homo sapien
16	2671	30.6	1905	1 LRP4_MOUSE	Q8vi56 mus musculus
17	2654	30.4	527	2 Q77501	Q77501 oryctolagus
18	2266.5	25.9	1768	2 Q7QEK9	Q7qek9 anopheles g
19	2228	25.5	2009	2 Q9VXN0	Q9vxm0 drosophila
20	2193	25.1	4544	1 LRP1_HUMAN	Q07954 homo sapien
21	2180	24.9	4543	1 LRP1_CHICK	Q98157 gallus gall
22	2165	24.8	4545	2 Q91ZK7	Q91zk7 mus musculus
23	2165	24.8	4545	2 Q920T4	Q920t4 mus musculus
24	2113	24.2	4545	2 Q61291	Q61291 mus musculus
25	2018	23.1	1731	2 Q8WY30	Q8wy30 homo sapien
26	2008	23.0	4599	1 LRP1B_HUMAN	Q9nzzr2 homo sapien
27	1997	22.8	4599	1 LRP1B_MOUSE	Q9i1i8 mus musculus
28	1876.5	21.5	4655	1 LRP2_HUMAN	Q98164 homo sapien
29	1876.5	21.5	4655	2 Q7Z5C0	Q7z5c0 homo sapien
30	1872.5	21.4	4655	2 Q7Z5C1	Q7z5c1 homo sapien
31	1856	21.2	4660	1 LRP2_RAT	P98158 rattus norv

32	1779.5	20.4	4569	2 Q7PS35	Q7ps35 anopheles g
33	1762.5	20.2	4699	2 Q9V383	Q9v383 drosophila
34	1759.5	20.1	4547	2 Q9W343	Q9w343 drosophila
35	1708	19.5	4569	2 Q7PV66	Q7pv66 anopheles g
36	1458	16.7	4753	1 LRP_CAEEL	Q04833 caenorhabdi
37	1044.5	12.0	1581	2 Q73809	Q73809 fugu rubrip
38	1024	11.7	252	2 Q9NSY4	Q9nsy4 homo sapien
39	980.5	11.2	1859	2 Q7PS28	Q7ps28 anopheles g
40	969	11.1	1809	2 Q8MP02	Q8mp02 periplaneta
41	914.5	10.5	1847	2 Q76952	Q76952 aedes aegypt
42	909.5	10.4	1984	1 VL_DROME	P98163 drosophila
43	902	10.3	1650	2 Q9QVT6	Q9qvt6 rattus sp.
44	876.5	10.0	881	2 Q8WY31	Q8wy31 homo sapien
45	859	9.8	1537	2 Q8WY29	Q8wy29 homo sapien

ALIGNMENTS

RESULT 1
QUES7 PRELIMINARY; PRT; 1615 AA.
AC QUES7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE LDL receptor member LR3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9908902; PubMed=9790987; DOI=10.1006/bbrc.1998.9545;
RA Dong Y., Lathrop W., Weaver D., Qiu Q., Cini J., Bertolini D.,
RA Chen D.;
RT "Molecular cloning and characterization of LR3, a novel LDL receptor
family protein with mitogenic activity.";
RL Biochem. Biophys. Res. Commun. 251:784-790(1998).
DR EMBL; AF077820; AAC72791.1; -.
DR HSP; P98162; LURP.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00057; Ldl_recept_a; 3.
DR Pfam; PF00058; Ldl_recept_b; 19.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00181; EGF; 4.
DR SMART; SM00192; LDLa; 3.
DR SMART; SM00135; LY; 20.
DR PROSITE; PS01209; LDLRA_1; 3.
DR PROSITE; PS50068; LDLRA_2; 3.
KW Receptor.
SQ SEQUENCE 1615 AA; 179143 MW; 8BA25D07F51E02CA CRC64;

Query Match 100.0%; Score 8740; DB 2; Length 1615;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 1615; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MEAAPGPPWPLLLLLLLLLLCCCPAPAAASPLLLFANRRDVLVDAGGKLESTIVVS	60
Db	1	MEAAPGPPWPLLLLLLLLLLCCCPAPAAASPLLLFANRRDVLVDAGGKLESTIVVS	60
Qy	61	GLEDAADVDFQSKGAVYWTVDVSEAIKQTYLNTGAAVQNVVIGSLVSPDGLACDWGK	120
Db	61	GLEDAADVDFQSKGAVYWTVDVSEAIKQTYLNTGAAVQNVVIGSLVSPDGLACDWGK	120
Qy	121	KLYWTDSETNRIEYANLNGTSRKVLPWQDLDPQRAIALDPAHGYMYTWDGCTPRIERAG	180
Db	121	KLYWTDSETNRIEYANLNGTSRKVLPWQDLDPQRAIALDPAHGYMYTWDGCTPRIERAG	180

Db 121 KLYWTDSETRIEVANLNGTSRKVLFWQDLQDRAIALDPAGHYMYWTDWGTGPRIERAG 180
Qy 181 MDGSTRKLIIVSDIYWPNGLTITDLBEOKLYWADAKLSFTHRANLDGSRQKVVESGLTHP 240
Db 181 MDGSTRKLIIVSDIYWPNGLTITDLBEOKLYWADAKLSFTHRANLDGSRQKVVESGLTHP 240
Qy 241 FALTLSGDTLWYTDWQTSIHACNKRITGSKREIISALYSMPDIOVLQOERQPPFHTTCE 300
Db 241 FALTLSGDTLWYTDWQTSIHACNKRITGSKREIISALYSMPDIOVLQOERQPPFHTTCE 300
Qy 301 EDNGCCHLCLLSPSEPFYTCACPTGVQLQDNGRTCKAGAEVLLARRTDLRLISLDT 360
Db 301 EDNGCCHLCLLSPSEPFYTCACPTGVQLQDNGRTCKAGAEVLLARRTDLRLISLDT 360
Qy 361 DFTDVLQVDDIRHAIADIDPLEGYVYVTDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
Db 361 DFTDVLQVDDIRHAIADIDPLEGYVYVTDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
Qy 421 IAVDWVARNLYWTDGTDRIEVTRLNGTSRKILVSEDLDEPRAIALHPVMGLMYWTDWGE 480
Db 421 IAVDWVARNLYWTDGTDRIEVTRLNGTSRKILVSEDLDEPRAIALHPVMGLMYWTDWGE 480
Qy 481 NPKIECANLDGQERRVLVNASLGNPGLALDLQEGKLYWDAKTDKIEVINVDGTKRRTL 540
Db 481 NPKIECANLDGQERRVLVNASLGNPGLALDLQEGKLYWDAKTDKIEVINVDGTKRRTL 540
Qy 541 LEDKLPHIFGFTLLGDFIYWTDWQRRSIERVHKYKASRDVLIIDQLPDLMLKAVNVAKV 600
Db 541 LEDKLPHIFGFTLLGDFIYWTDWQRRSIERVHKYKASRDVLIIDQLPDLMLKAVNVAKV 600
Qy 601 GTNPCADNRGCGSHLCFFTPHATRCGCPGLELLSDMKTCIVPEAFVFTSRAAHLISL 660
Db 601 GTNPCADNRGCGSHLCFFTPHATRCGCPGLELLSDMKTCIVPEAFVFTSRAAHLISL 660
Qy 661 ETNNNDVAIPUTGVKEASALDFVSNHHIYWTDVSLKTIISRAFMNGSSVERVVFGLDYP 720
Db 661 ETNNNDVAIPUTGVKEASALDFVSNHHIYWTDVSLKTIISRAFMNGSSVERVVFGLDYP 720
Qy 721 EGMADVMMGNLYWADTGTNRIEVARLDGQFVRLVMDLONPSLALDPTKGYIYWTW 780
Db 721 EGMADVMMGNLYWADTGTNRIEVARLDGQFVRLVMDLONPSLALDPTKGYIYWTW 780
Qy 781 CGKPRIVRAFMNDGNTCMVLDKVGANDLTIDYADQRLYWDLTDTNMISSNMLGQERVV 840
Db 781 CGKPRIVRAFMNDGNTCMVLDKVGANDLTIDYADQRLYWDLTDTNMISSNMLGQERVV 840
Qy 841 IADDLPHFPGLTQYSDIYIYWTDMNLHSTIERADKTSGRNRTLIOGHLPFVMDILVFHSSRQ 900
Db 841 IADDLPHFPGLTQYSDIYIYWTDMNLHSTIERADKTSGRNRTLIOGHLPFVMDILVFHSSRQ 900
Qy 901 DGLNDCMHNNGCCOLCLAIIPGHRCCASHYTLDPGSRNCSPPTFLFSQKSAISMI 960
Db 901 DGLNDCMHNNGCCOLCLAIIPGHRCCASHYTLDPGSRNCSPPTFLFSQKSAISMI 960
Qy 961 PDDQHSPLILPLHGLRNVAKIDYDPLDKFYIYWDGQNIKRAKDDGTQPPVLTSLSQGG 1020
Db 961 PDDQHSPLILPLHGLRNVAKIDYDPLDKFYIYWDGQNIKRAKDDGTQPPVLTSLSQGG 1020
Qy 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEMGVVLRGDRPKPRAIVVNAERG 1080
Db 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEMGVVLRGDRPKPRAIVVNAERG 1080
Qy 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWADLAKRIESCD 1140
Db 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWADLAKRIESCD 1140
Qy 1141 LSGANRLTLEDANIVQPLGLTILGKHLVYIDRQOQMIEVERKTTGDKRTRIQGRVAHLTG 1200
Db 1141 LSGANRLTLEDANIVQPLGLTILGKHLVYIDRQOQMIEVERKTTGDKRTRIQGRVAHLTG 1200
Qy 1201 IHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQNLITCGEPPTCS 1260
Db 1201 IHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQNLITCGEPPTCS 1260

RESULT 2

LRP5 HUMAN

ID LRP5 HUMAN STANDARD; PRT; 1615 AA.

AC 075197; 096TD6; Q9UP66;
DT 05-JUL-2004 (Rel. 44, Created)

DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Low-density lipoprotein receptor-related protein 5 precursor.

GN Name=LRP5; Synonyms=LRP7;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Osteoblast;

RX MEDLINE=98382578; PubMed=9714764; DOI=10.1016/S0378-1119(98)00311-4;

RA Hey P.J.; Twells R.C.J.; Phillips M.S.; Nakagawa Y.; Brown S.D.,

RA Kawaguchi Y.; Cox R.; Xie G.; Dugan V.; Hammond H.; Metzker M.D.,

RA Todd J.A.; Hess J.F.;

RT "Cloning of a novel member of the low-density lipoprotein receptor

family."

RL Gene 216:103-111(1998).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=21295044; PubMed=11401438; DOI=10.1006/geno.2000.6492;

RA Twells R.C.J.; Metzker M.D.; Brown S.D.; Cox R.; Garey C.; Hammond H.,

RA Hey P.J.; Levy E.; Nakagawa Y.; Phillips M.S.; Todd J.A.; Hess J.F.;

RT "The sequence and gene characterization of a 400-kb candidate region

for IDDM4 on chromosome 11q13."

RL Genomics 72:231-242(2001).

RN [3]

RP SEQUENCE FROM N.A.

RX PubMed=12509515; DOI=10.1073/pnas.0133792100;

RA Fujino T.; Asaba H.; Kang M.J.; Ikeda Y.; Sone H.; Takada S.,

RA Kim D.H.; Ioka R.X.; Ono M.; Tomoyori H.; Okubo M.; Murase T.,

RA Kamataki A.; Yamamoto J.; Magoori K.; Takahashi S.; Miyamoto Y.,

RA Oishi H.; Nose M.; Okazaki M.; Usui K.; Imaizumi K.; Yanagisawa M.,

RA Sakai J.; Yamamoto T.T.;

RT "Low-density lipoprotein receptor-related protein 5 (LRP5) is

essential for normal cholesterol metabolism and glucose-induced

insulin secretion."

RL Proc. Natl. Acad. Sci. U.S.A. 100:229-234(2003).

RN [4]

RP FUNCTION, PHOSPHORYLATION, AND INTERACTION WITH AXIN.

RX PubMed=14731402; DOI=10.1016/S1097-2765(03)00484-2;

RA Tamai K.; Zeng X.; Liu C.; Zhang X.; Harada Y.; Chang Z.; He X.;

RT "A mechanism for Wnt coreceptor activation.";
 RL Mol. Cell 13:149-156(2004).
 RN [5]
 RP REVIEW
 RX PubMed=15084453; DOI=10.1342/dev.01117;
 RA He X., Semenov M., Tamai K., Zeng X.;
 RL "LDL receptor-related proteins 5 and 6 in Wnt/beta-catenin signaling:
 RT arrows point the way.";
 RN Development 131:1663-1677(2004).
 RL [6]
 RP VARIANTS MET-173; HIS-1168 AND GLY-1361.
 RX PubMed=15024691; DOI=10.1086/383202;
 RA Toomes C., Bottomley H.M., Jackson R.M., Towns K.V., Scott S.,
 RA Mackey D.A., Craig J.E., Jiang L., Yang Z., Trembath R., Woodruff G.,
 RA Gregory-Evans C.Y., Gregory-Evans K., Parker M.J., Black G.C.,
 RA Downey L.M., Zhang K., Inglehearn C.F.;
 RL "Mutations in LRP5 or FZD4 underlie the common familial exudative
 RT vitreoretinopathy locus on chromosome 11q.";
 RL Am. J. Hum. Genet. 74:721-730(2004).
 CC -1- FUNCTION: Involved in the Wnt/beta catenin signaling pathway,
 CC probably by acting as a coreceptor together with Frizzled for Wnt.
 CC -1- SUBUNIT: Interacts with different Wnt/Frizzled complexes.
 CC interacts with axin.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: Widely expressed, with the highest level of
 CC expression in the liver.
 CC -1- PMW: Phosphorylation of the PPPSP motif creates an inducible
 CC docking site for axin.
 CC -1- POLYMORPHISM: The Met-17; His-1168 and Gly-1361 variants were
 CC found in family with FEVER [MIM:605750]. However, asymptomatic
 CC individuals were also found to have the mutations. It is thus
 CC unknown whether they are disease-causing mutations.
 CC -1- DISEASE: Defects in LRP5 may be a cause of familial exudative
 CC vitreoretinopathy (FEVR) [MIM:605750]. FEVR is a disorder of the
 CC retinal vasculature characterized by an abrupt cessation of the
 CC growth of peripheral retinal capillaries. Autosomal dominant
 CC inheritance is the most common mode of segregation, although
 CC recessive and X-linked pedigrees are also seen. The phenotype is
 CC variable. In severe cases, the formation of fibrovascular vitreous
 CC membranes that cause retinal traction may result in the
 CC displacement of the macula, the presence of retinal folds, and/or
 CC the detachment of the retina, any of which may cause a decrease in
 CC visual acuity sufficient to be labeled legal blindness. Patients
 CC with mild forms of the disease are asymptomatic, and their only
 CC disease-related abnormality is an arc of avascular retina in the
 CC extreme temporal periphery.
 CC -1- DISEASE: Defects in LRP5 may be a cause of osteoporosis
 CC pseudoglioma syndrome (OPPG) [MIM:259770]. OPPG is a recessive
 CC disorder characterized by very low bone mass and blindness.
 CC Individually with OPPG are prone to develop bone fractures and
 CC deformations and have various eye abnormalities, including
 CC phthisis bulbi, retinal detachments, falciform folds or persistent
 CC vitreal vasculature.
 CC -1- SIMILARITY: Belongs to the LDLR family.
 CC -1- SIMILARITY: Contains 4 EGF-like domains.
 CC -1- SIMILARITY: Contains 3 LDL-receptor class A domains.
 CC -1- SIMILARITY: Contains 19 LDL-receptor class B domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/ebis-sib.ch/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AF064548; AA36467.1; --
 CC EMBL; AF283321; AA52433.1; --
 CC EMBL; AF283320; AA52433.1; JOINED.
 CC EMBL; AB017498; BA33051.1; --
 CC PIR; J03372; J03372.
 CC HSSP; P98162; 1JRF.
 CC Genew; HGNC:6697; LRP5.

DR MIM; 603506; --
 DR MIM; 605750; --
 DR MIM; 259770; --
 DR GO; GO:0016021; C: integral to membrane; TAS.
 DR GO; GO:0006629; P: lipid metabolism; TAS.
 DR GO; GO:0008284; P: positive regulation of cell proliferation; TAS.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR008210; EGF.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR InterPro; IPR000033; LDL_receptor_rep.
 DR InterPro; IPR011042; TolB_C.
 DR Pfam; PF00008; EGF_4.
 DR Pfam; PF00057; Ldl_recept_a; 3.
 DR Pfam; PF00058; Ldl_recept_b; 19.
 DR PRINTS; PR00261; LDLRECEPTOR.
 DR SMART; SM00181; EGF; 4.
 DR SMART; SM00192; LDLA; 3.
 DR SMART; SM00135; LY; 20.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; FALSE_NEG.
 DR PROSITE; PS00026; EGF_3; FALSE_NEG.
 DR PROSITE; PS01187; EGF_CA; FALSE_NEG.
 DR PROSITE; PS01209; LDLRA_1; 3.
 DR PROSITE; PS00068; LDLRA_2; 3.
 KW Developmental protein; EGF-like domain; Endocytosis; Glycoprotein;
 KW Phosphorylation; Polymorphism; Receptor; Repeat; Signal;
 KW Transmembrane; Wnt signaling pathway.
 FT SIGNAL 1 31
 FT CHAIN 32 1615
 FT DOMAIN 32 1384
 FT TRANSMEM 1385 1407
 FT DOMAIN 1408 1615
 FT DOMAIN 75 118
 FT DOMAIN 120 161
 FT DOMAIN 163 205
 FT DOMAIN 207 246
 FT DOMAIN 248 289
 FT DOMAIN 295 337
 FT DOMAIN 385 426
 FT DOMAIN 428 469
 FT DOMAIN 471 513
 FT DOMAIN 515 558
 FT DOMAIN 601 641
 FT DOMAIN 687 728
 FT DOMAIN 730 771
 FT DOMAIN 773 814
 FT DOMAIN 816 853
 FT DOMAIN 856 897
 FT DOMAIN 902 942
 FT DOMAIN 989 1034
 FT DOMAIN 1036 1077
 FT DOMAIN 1079 1122
 FT DOMAIN 1124 1162
 FT DOMAIN 1165 1207
 FT DOMAIN 1213 1254
 FT DOMAIN 1258 1296
 FT DOMAIN 1297 1333
 FT DOMAIN 1335 1371
 FT DOMAIN 1495 1610
 FT REPEAT 78 81
 FT REPEAT 123 126
 FT REPEAT 166 169
 FT REPEAT 251 254
 FT REPEAT 388 391
 FT REPEAT 431 434
 FT REPEAT 474 477
 FT REPEAT 559 562
 FT REPEAT 690 693
 FT REPEAT 819 822
 FT REPEAT 859 862

RL Gene 216:103-111 (1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND ALTERNATIVE SPLICING.
RX TISSUE=Liver;
RC MEDLINE=99168901; PubMed=10049586; DOI=10.1006/geno.1998.5688;
RA Chen D., Lathrop W., Dong Y.;
RT "Molecular cloning of mouse Lrp7(lr3) cDNA and chromosomal mapping of
RL orthologous genes in mouse and human.";
RN Genomics 55:314-321(1999).
[3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX TISSUE=Breast tumor;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Usdin T.B., Bonaldo M.F., Casavant T.L., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Greenwood J., Schmutz J., Myers R.M.,
RA Butterfield A.C., Krzyzanski M.I., Skalska U., Smalls D.B.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- FUNCTION: Involved in the Wnt/beta catenin signaling pathway, for Wnt
CC probably by acting as a coreceptor together with Frizzled for Wnt
CC (By similarity).
CC -1- SUBUNIT: Interacts with different Wnt/Frizzled complexes.
CC Interacts with axin (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=1;
CC Comment=A number of isoforms are produced;
CC Name=1;
CC IsoId=Q91VN0-1; Sequence=Displayed;
CC -1- TISSUE SPECIFICITY: Widely expressed, with the highest expression
CC levels in liver, heart, and lung and the lowest levels in brain
CC and spleen.
CC -1- DEVELOPMENTAL STAGE: Expressed before or on embryonic day 7.
CC -1- PTM: Phosphorylation of the PPPSP motif creates an inducible
CC docking site for axin (By similarity).
CC -1- SIMILARITY: Belongs to the LDLR family.
CC -1- SIMILARITY: Contains 4 EGF-like domains.
CC -1- SIMILARITY: Contains 19 LDL-receptor class A domains.
CC -1- SIMILARITY: Contains 19 LDL-receptor class B domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF064984; AAC36468.1; -;
CC EMBL; AF077847; AAC70183.1; -;
CC EMBL; BC011374; AAH11374.1; -;
CC HSSP; P01130; 1AJJ.
CC MGD; MGI:1278315; Lrp5.
CC InterPro; IPR006209; EGF like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR002172; LDL_receptor A.
CC InterPro; IPR000033; Ldl_receptor_rep.
CC Pfam; PF00008; EGF; 4.
DR DR Pfam; PF00057; Ldl_recept_a; 3.
DR DR Pfam; PF00058; Ldl_recept_b; 19.
DR DR PRINTS; PR00261; LDLRECEPTOR.
DR DR SMART; SM00181; EGF; 4.
DR DR SMART; SM00192; LDLa; 3.
DR DR SMART; SM00135; LY; 20.
DR DR PROSITE; PS00022; EGF_1; FALSE NEG.
DR DR PROSITE; PS01186; EGF_2; FALSE NEG.
DR DR PROSITE; PS00026; EGF_3; FALSE NEG.
DR DR PROSITE; PS01187; EGF_CA; FALSE NEG.
DR DR PROSITE; PS01209; LDLRA_1; 3.
DR DR PROSITE; PS00068; LDLRA_2; 3.
DR DR Alternative splicing; Developmental protein; EGF-like domain;
KW Endocytosis; Glycoprotein; Phosphorylation; Receptor; Repeat; Signal;
KW Transmembrane; Wnt signaling pathway.
FT SIGNAL 1 30 Potential.
FT CHAIN 31 1614 Low-density lipoprotein receptor-related
FT protein 5.
FT DOMAIN 31 1383 Extracellular (Potential).
FT TRANSMEM 1384 1406 Potential.
FT DOMAIN 1407 1614 Cytoplasmic (Potential).
FT DOMAIN 74 117 LDL-receptor class B 1.
FT DOMAIN 119 160 LDL-receptor class B 2.
FT DOMAIN 162 204 LDL-receptor class B 3.
FT DOMAIN 206 245 LDL-receptor class B 4.
FT DOMAIN 247 288 LDL-receptor class B 5.
FT DOMAIN 294 336 EGF-like 1.
FT DOMAIN 384 425 LDL-receptor class B 6.
FT DOMAIN 427 468 LDL-receptor class B 7.
FT DOMAIN 470 512 LDL-receptor class B 8.
FT DOMAIN 514 557 LDL-receptor class B 9.
FT DOMAIN 600 640 EGF-like 2.
FT DOMAIN 686 727 LDL-receptor class B 10.
FT DOMAIN 729 770 LDL-receptor class B 11.
FT DOMAIN 772 813 LDL-receptor class B 12.
FT DOMAIN 815 852 LDL-receptor class B 13.
FT DOMAIN 855 896 LDL-receptor class B 14.
FT DOMAIN 901 941 EGF-like 3.
FT DOMAIN 988 1033 LDL-receptor class B 15.
FT DOMAIN 1035 1076 LDL-receptor class B 16.
FT DOMAIN 1078 1121 LDL-receptor class B 17.
FT DOMAIN 1123 1164 LDL-receptor class B 18.
FT DOMAIN 1165 1206 LDL-receptor class B 19.
FT DOMAIN 1212 1253 EGF-like 4.
FT DOMAIN 1257 1295 LDL-receptor class A 1.
FT DOMAIN 1296 1332 LDL-receptor class A 2.
FT DOMAIN 1334 1370 LDL-receptor class A 3.
FT SITE 1499 1503 PPPSP motif.
FT SITE 1605 1609 By similarity.
FT DISULFID 298 309 By similarity.
FT DISULFID 305 320 By similarity.
FT DISULFID 322 335 By similarity.
FT DISULFID 604 615 By similarity.
FT DISULFID 611 624 By similarity.
FT DISULFID 626 639 By similarity.
FT DISULFID 905 916 By similarity.
FT DISULFID 912 925 By similarity.
FT DISULFID 927 940 By similarity.
FT DISULFID 1216 1227 By similarity.
FT DISULFID 1223 1237 By similarity.
FT DISULFID 1239 1252 By similarity.
FT DISULFID 1258 1272 By similarity.
FT DISULFID 1265 1285 By similarity.
FT DISULFID 1279 1294 By similarity.
FT DISULFID 1297 1309 By similarity.
FT DISULFID 1304 1322 By similarity.
FT DISULFID 1316 1331 By similarity.
FT DISULFID 1335 1347 By similarity.
FT DISULFID 1342 1360 By similarity.
FT DISULFID 1354 1369 By similarity.
FT CARBOHYD 92 92 N-linked (GLCNAC. . .) (Potential).
FT CARBOHYD 137 137 N-linked (GLCNAC. . .) (Potential).
FT CARBOHYD 445 445 N-linked (GLCNAC. . .) (Potential).

FT CARBOHYD 498 498 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 704 704 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 877 877 N-linked (GlcNAc...) (Potential).
FT CONFLICT 220 220 R -> H (in Ref. 2).
FT CONFLICT 1520 1520 I -> S (in Ref. 3).
FT CONFLICT 1553 1553 I -> T (in Ref. 3).
SQ SEQUENCE 1614 AA; 178896 MW; 911EB288471752C5 CRC64;

Query Match 94.5%; Score 8259.5; DB 1; Length 1614;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1519; Conservative 41; Mismatches 51; Indels 7; Gaps 2;

QY 1 MEAP-...PGPEWPLLLLLLALCGCPAPAAASPLLFANRRDRLVLDAGGVKLSSTI 57
DB 1 METAPTRAPPPLPPPLLLLVYCSL----VPAASPLLFANRRDRLVLDAGGVKLSSTI 56
QY 58 VVSGLEDAADVDFQSKGAVYTDVSEBAIKQTYLNQTGAAYVNVISGLVSPDGLACDW 117
DB 57 VASGLEDAADVDFQSKGAVYTDVSEBAIKQTYLNQTGAAYVNVISGLVSPDGLACDW 116
QY 118 VGKLYWTDSETNRIEVANLNGTSRKVLFWODLQDPAIRALDPAHYMYTWDGSETPRIE 177
DB 117 VGKLYWTDSETNRIEVANLNGTSRKVLFWODLQDPAIRALDPAHYMYTWDGSEAPRIE 176
QY 178 RAGMDGSTRKIIIVDSDIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPQKVVESGL 237
DB 177 RAGMDGSTRKIIIVDSDIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPQKVVESGL 236
QY 238 THPPALTLGDTLWTDQWTSIIHACNKTGGKKEILSALYSPMDIOVLQSOERPPFHT 297
DB 237 THPPALTLGDTLWTDQWTSIIHACNKTGGKKEILSALYSPMDIOVLQSOERPPFHT 296
QY 298 RCEBDGSCSHLCLLSSEPEYTCACPTGVLOQNGRTCKAGAEVLLIARTRILSL 357
DB 297 PCEBDGSCSHLCLLSSEPEYTCACPTGVLOQNGRTCKAGAEVLLIARTRILSL 356
QY 358 DTPDFTDVLQVDDIRHAIAIDYDPLEGVYVYTDDEVAIRRAYLDGSGAOTLVNTEIND 417
DB 357 DTPDFTDVLQVDDIRHAIAIDYDPLEGVYVYTDDEVAIRRAYLDGSGAOTLVNTEIND 416
QY 418 PDGIAVDWAAVNLWTDGTDRIEVTRNLGTSRKILVSEDLDEPRAIALHPVGMVMTWD 477
DB 417 PDGIAVDWAAVNLWTDGTDRIEVTRNLGTSRKILVSEDLDEPRAIALHPVGMVMTWD 476
QY 478 WGENPKTECANLQGERVLVNASIGWPNGLALDLOEGKLYWDAKTDKIEVINVDGTR 537
DB 477 WGENPKTECANLQGERVRLVNTSLGPNGLALDLOEGKLYWDAKTDKIEVINVDGTR 536
QY 538 RTLEDKLPHPFGFTLLGDFIYTDWQRRSIRVHKVKAARDVLIIDQLPDLMLGKAVNVA 597
DB 537 KTLLEDKLPHPFGFTLLGDFIYTDWQRRSIRVHKVKAARDVLIIDQLPDLMLGKAVNVA 596
QY 598 KVGVTNCPADRGCGSHLCFTPTPHATRCGCPIGLELLSDMKTCIVPRAFLVFTSRAIHR 657
DB 597 KVGVTNCPADRGCGSHLCFTPTPHATRCGCPIGLELLSDMKTCIPEAFVFTSRAIHR 656
QY 658 ISLETNNDVAILPTGVKEASALDPVDSNNHIIYTDVSLKTIISRAFWNGSSVEHVVERGL 717
DB 657 ISLETNNDVAILPTGVKEASALDPVDSNNHIIYTDVSLKTIISRAFWNGSSVEHVVERGL 716
QY 718 DYPEGMAVDMWGMKNLYWADTGTNRIEVARLDGQFQVLRDLNPRSLALDPTKGIYV 777
DB 717 DYPEGMAVDMWGMKNLYWADTGTNRIEVARLDGQFQVLRDLNPRSLALDPTKGIYV 776
QY 778 TEWCGKPRIVAFWDGNTCMFLVDKVGANDLTIDYADQRLYWTDLDTNMISSNMLGQE 837
DB 777 TEWCGKPRIVAFWDGNTCMFLVDKVGANDLTIDYADQRLYWTDLDTNMISSNMLGQE 836
QY 838 RVVIADLLPHFPFGITQYSDIYTWDMNLHSLERADKTSGRNRTLIQGHLDVMDILVFS 897
DB 837 RVVIADLLPHFPFGITQYSDIYTWDMNLHSLERADKTSGRNRTLIQGHLDVMDILVFS 896
QY 898 SRQGLNDCHMNGCCGQOLCLAIPEGHRCGCASHYTLDPSSRNCSPPTFTLLFSQKSAIS 957

DB 897 SRQGLNDCHMNGCCGQOLCLAIPEGHRCGCASHYTLDPSSRNCSPPTFTLLFSQKSAIS 956
QY 958 RMIPDDQHSPLILPLHGLRNKKAIDYDPLDKFIYWDGQRNIKRAKDDGTQPPVLTSL 1017
DB 957 RMIPDDQHSPLILPLHGLRNKKAIDYDPLDKFIYWDGQRNIKRAKDDGTQPPVLTSL 1016
QY 1018 QGQNPDRQPHDLSIDYSRITLFWTCEATNTINVRHLSGEAMGVLRGDRKPRRAIVNAE 1077
DB 1017 QLSFDRQPHDLSIDYSRITLFWTCEATNTINVRHLSGEAMGVLRGDRKPRRAIVNAE 1076
QY 1078 RGYLYFTNMQDRAAKIERAALDGTREVLFTGLIRPVVALVVDNLGKLFWDADLKR 1137
DB 1077 RGYLYFTNMQDRAAKIERAALDGTREVLFTGLIRPVVALVVDNLGKLFWDADLKR 1136
QY 1138 SCDLSGANRLTLEDANIVQPLGLTILGKHLWIDRQQQMIERVEKTTGDKTRIQGRAH 1197
DB 1137 SCDLSGANRLTLEDANIVQPLGLTILGKHLWIDRQQQMIERVEKTTGDKTRIQGRAH 1196
QY 1198 LTGHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSCPVHLVLLQNLTCGEP 1257
DB 1197 LTGHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSCPVHLVLLQNLTCGEP 1256
QY 1258 TCSPDQACATGEIDCIPGAWRCDFPECDQSBEGCPVCSAAQPCARGOCVDLRRC 1317
DB 1257 TCSPDQACATGEIDCIPGAWRCDFPECDQSBEGCPVCSAAQPCARGOCVDLRRC 1316
QY 1318 DGEADQDRSDEADCAICLPNQPRCAGQCCLIKQCCDSDPDICIDGSDLMCEITKPPS 1377
DB 1317 DGEADQDRSDEADCAICLPNQPRCAGQCCLIKQCCDSDPDICIDGSDLMCEITKPPS 1376
QY 1378 DDPASHASATGPVIGIILSLFVMGVYVFCORVVCORVYAGANGPPHYVSGTHVPLNF 1437
DB 1377 DDPASHASATGPVIGIILSLFVMGVYVFCORVVCORVYAGANGPPHYVSGTHVPLNF 1436
QY 1438 IAPGSGHQGPPTGIACGKSMSSVSLMGGRCVPLYDRNHVHTGASSSSSSSTKATLYPPI 1497
DB 1437 IAPGSGHQGPPTGIACGKSMSSVSLMGGRCVPLYDRNHVHTGASSSSSSSTKATLYPPI 1496
QY 1498 LNPPSPATDPSLVNMDYSSNIPATAPRVPYIIRGMAPPPTPCSTDVCDSDYSASRW 1557
DB 1497 LNPPSPATDPSLVNMDYSSNIPATAPRVPYIIRGMAPPPTPCSTDVCDSDYSASRW 1556
QY 1558 KASKYVLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
DB 1557 KASKYVLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1614

RESULT 4
Q8AYF1 PRELIMINARY; PRT; 1605 AA.
AC Q8AYF1
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Lipoprotein receptor-related protein 5.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22194715; PubMed=12204281; DOI=10.1016/S0925-4773(02)00205-8;
RA Houston D.W., Wylie C.;
RT "Cloning and expression of Xenopus Lrp5 and Lrp6 genes.";
RL Mech. Dev. 117:337-342(2002).
DR EMBL; AF508960; AA09806.1; -.
DR HSSP; P01130; 1D2J.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.

"Isolation and Characterization of LRP6, a novel member of the low density lipoprotein receptor gene family.";
 Biochem. Biophys. Res. Commun. 248:879-888(1998).
 [2]
 FUNCTION, AND DOMAINS REQUIRED FOR INTERACTION WITH DKK1 AND DKK2.
 PubMed=11357136; DOI=10.1038/35077108;
 RA Mac B., Wu W., Li Y., Hoppe D., Stanek P., Glinka A., Niehrs C.;
 "LDL-receptor-related protein 6 is a receptor for Dickkopf proteins.";
 Nature 411:321-325(2001).
 [3]
 INTERACTIONS WITH FZD5; DKK1 AND DKK2.
 PubMed=12857724; DOI=10.1074/jbc.M300191200;
 RA Caricasole A., Ferraro T., Iacovelli L., Barletta E., Caruso A.,
 RA Melchiorri D., Terstappen G.C., Nicoletti F.;
 "Functional characterization of WNT7A signaling in PC12 cells: by
 interaction with A FZD5 x LRP6 receptor complex and modulation by
 Dickkopf proteins.";
 J. Biol. Chem. 278:37024-37031(2003).
 [4]
 REVIEW
 PubMed=15084453; DOI=10.1242/dev.01117;
 RA He X., Semenov M., Tamai K., Zeng X.;
 "LDL receptor-related proteins 5 and 6 in Wnt/[beta]-catenin
 signaling: Arrows point the way.";
 Development 131:1663-1677(2004).
 CC -1- FUNCTION: Essential for the Wnt/beta catenin signaling pathway,
 probably by acting as a coreceptor together with Frizzled for Wnt.
 CC Specific high-affinity receptor for DKK1 and DKK2, but not DKK3.
 CC The interaction with DKK1 blocks LRP6-mediated Wnt/beta catenin
 CC signaling via LRP6 removal via Kremen proteins-mediated
 CC endocytosis.
 CC -1- SUBUNIT: Interacts with FZD5. Essential component of the Wnt7
 CC receptor complex. Wnt7A interacts with the LRP6/FZD5 complex. This
 CC interaction is antagonized by DKK1 and DKK3.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: Widely co-expressed with LRP5 during
 CC embryogenesis and in adult tissues.
 CC -1- DOMAIN: The WNT-EGF-like domains 1 and 2 are required for the
 CC interaction with Wnt-frizzled complex. The WNT-EGF-like domains 3
 CC and 4 are required for the interaction with DKK1.
 CC -1- SIMILARITY: Belongs to the LDLR family.
 CC -1- SIMILARITY: Contains 4 EGF-like domains.
 CC -1- SIMILARITY: Contains 3 LDL-receptor class A domains.
 CC -1- SIMILARITY: Contains 20 LDL-receptor class B domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF074264; AAC33006.1; -;
 CC PIR; J0272; J0272.
 CC HSP; P01130; 1AJJ.
 CC Genew; HGNC:6698; LRP6.
 CC MIM; 603507; -;
 CC GO; GO:0016021; C: integral to membrane; TAS.
 CC GO; GO:0005886; C: plasma membrane; TAS.
 CC GO; GO:0005515; F: protein binding; TAS.
 CC InterPro; IPR000742; EGF 2.
 CC InterPro; IPR006209; EGF-like.
 CC InterPro; IPR006210; IEGF.
 CC InterPro; IPR002172; LDL_receptor A.
 CC InterPro; IPR000033; Ldl_receptor_rep.
 CC Pfam; PF00008; EGF; 4.
 CC Pfam; PF00057; Ldl_recept_a; 3.
 CC Pfam; PF00058; Ldl_recept_b; 20.
 CC PRINTS; PR00261; LDLRECEPTOR.
 CC SMART; SM00181; EGF; 4.
 CC SMART; SM00192; LDLa; 3.
 CC SMART; SM00135; LY; 20.

DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00026; EGF_3; FALSE_NEG.
 DR PROSITE; PS01209; LDLRA 1; 3.
 DR PROSITE; PS00069; LDLRA 2; 3.
 KW Developmental protein; EGF-like domain; Endocytosis; Glycoprotein;
 KW Receptor; Repeat; Signal; Transmembrane; Wnt signaling pathway.
 FT SIGNAL 1 19 Potential.
 FT CHAIN 20 1613 Low-density lipoprotein receptor-related protein 6.
 FT DOMAIN 20 1370 Extracellular (Potential).
 FT TRANSMEM 1371 1393 Potential.
 FT DOMAIN 1394 1613 Cytoplasmic (Potential).
 FT DOMAIN 63 105 LDL-receptor class B 1.
 FT DOMAIN 107 148 LDL-receptor class B 2.
 FT DOMAIN 150 192 LDL-receptor class B 3.
 FT DOMAIN 194 235 LDL-receptor class B 4.
 FT DOMAIN 237 276 LDL-receptor class B 5.
 FT DOMAIN 282 324 EGF-like 1.
 FT DOMAIN 372 413 LDL-receptor class B 6.
 FT DOMAIN 415 456 LDL-receptor class B 7.
 FT DOMAIN 458 500 LDL-receptor class B 8.
 FT DOMAIN 502 540 LDL-receptor class B 9.
 FT DOMAIN 543 586 LDL-receptor class B 10.
 FT DOMAIN 588 628 EGF-like 2.
 FT DOMAIN 674 715 LDL-receptor class B 11.
 FT DOMAIN 717 758 LDL-receptor class B 12.
 FT DOMAIN 760 801 LDL-receptor class B 13.
 FT DOMAIN 803 840 LDL-receptor class B 14.
 FT DOMAIN 843 884 LDL-receptor class B 15.
 FT DOMAIN 889 930 EGF-like 3.
 FT DOMAIN 977 1024 LDL-receptor class B 16.
 FT DOMAIN 1026 1067 LDL-receptor class B 17.
 FT DOMAIN 1069 1112 LDL-receptor class B 18.
 FT DOMAIN 1114 1155 LDL-receptor class B 19.
 FT DOMAIN 1157 1197 LDL-receptor class B 20.
 FT DOMAIN 1203 1244 EGF-like 4.
 FT DOMAIN 1248 1286 LDL-receptor class A 1.
 FT DOMAIN 1287 1323 LDL-receptor class A 2.
 FT DOMAIN 1325 1361 LDL-receptor class A 3.
 FT SITE 1487 1491 PPPSP motif.
 FT SITE 1604 1608 PPPSP motif.
 FT DISULFID 286 297 By similarity.
 FT DISULFID 293 308 By similarity.
 FT DISULFID 310 323 By similarity.
 FT DISULFID 592 603 By similarity.
 FT DISULFID 599 612 By similarity.
 FT DISULFID 614 627 By similarity.
 FT DISULFID 893 904 By similarity.
 FT DISULFID 900 914 By similarity.
 FT DISULFID 916 929 By similarity.
 FT DISULFID 1207 1218 By similarity.
 FT DISULFID 1214 1228 By similarity.
 FT DISULFID 1230 1243 By similarity.
 FT DISULFID 1249 1263 By similarity.
 FT DISULFID 1256 1276 By similarity.
 FT DISULFID 1270 1285 By similarity.
 FT DISULFID 1288 1300 By similarity.
 FT DISULFID 1295 1313 By similarity.
 FT DISULFID 1307 1322 By similarity.
 FT DISULFID 1326 1338 By similarity.
 FT DISULFID 1333 1351 By similarity.
 FT DISULFID 1345 1360 By similarity.
 FT CARBOHYD 42 42 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 81 81 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 281 281 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 433 433 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 486 486 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 592 592 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 859 859 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 865 865 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 926 926 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1039 1039 N-linked (GlcNAc...) (Potential).

SQ	SEQUENCE	1613 AA; 180442 MW; 4BC1141E395D8B5B CRC64;
Query Match	70.4%; Score 6153.5; DB 1; Length 1613;	
Best Local Similarity	70.2%; Pred. No. 0;	
Matches 1134; Conservative 205; Mismatches 256; Indels 21; Gaps 12;		
Qy	15 LLLLLLALCGCPAPAAASPLLLFANRRDRVLVDAGGVKLESTIVVSGLEDAADVPFQSK 74	
Db	4 VLRSLLACSFVLLRAA-PLLLYANRRDLRLVDATNGKENATIVVGGLEDAADVPFVFSH 62	
Qy	75 GAVYWTDSBEAIIKOTYLNOGAQVNVVIGSLVSPDGLACDWGKKLYWTDSTNRIEV 134	
Db	63 GLIYWSVSEBAIKRTFENKT-ESQNVVVVSGLLSPDGLACDWGKKLYWTDSTNRIEV 121	
Qy	135 ANLNGTSRKVLFWQDLQOPRAIALDPDAHGYMYWTDWGETPRIERAGMDGSTRKIIVDSDI 194	
Db	122 SNLGSRLKVLFWQDLQOPRAIALDPSSGFMYWTDWGEVPIKIERAGMDGSSRFIIINSEI 181	
Qy	195 YWPNGLTIDLEOKLYWADAKLSFIHRANLDGSPROKVVGEGLTHPFAALTUSGDTLIWTD 254	
Db	182 YWPNGLTIDYBEOKLYWADAKLFIHKSNDLGTNRQA VVKGSLPHFPFALTIFEDILYWTD 241	
Qy	255 WQTESIHACNKRCTGKGRKEIISALYSMDIQVLQOEROPFHTCEBNDGCGSCHLLISP 314	
Db	242 WSTHSIIACNKYTGELUREIHSIDIFSPMDIHAFSQOQPNATPCGIDNGCGSCHLCMSP 301	
Qy	315 SEPFYTCACPTGVOLQNGRTCKAGAEVILLARLTDLRRISLDTPOFTDIVLQVDDIRH 374	
Db	302 VKPFYQCACPTGVKLENGTKCKGATELLLLARLTDLRRISLDTPOFTDIVLQLEDIRH 361	
Qy	375 AIAIDYDPLEGYVYWTDEVAIRAYLDGSGAOTLVNTEINDPDGIADVWVARNLYWTD 434	
Db	362 AIAIDYDPEGYIYWTDEVAIRARSFTDGSQGFVYTAQIAHPDGIADVWVARNLYWTD 421	
Qy	435 TGTDRIEVTRLNGTSRKILVSEDLDEPRAIALHPVWGLMYWTDWGENPKIECANLDQER 494	
Db	422 TGTDRIEVTRLNGTMRKILVSEDLDEPRAIYVLDPMVGYMYWTDWGEIPKIERAALDGSDR 481	
Qy	495 RVLVNASLWPNGLALDLOEGLKLYWDAKTDKIEVINVDGTRKRTLLDKLPHIFGFTLL 554	
Db	482 VLVNTSLWPNGLALDYDESKIYWGDAKTDKIEVMWTDGTGRRVLVEDKIPHIFGFTLL 541	
Qy	555 GDFIYWTDWQRSSIERVHKVKSARDVIIDQLPDLMLKANVAVGNTNPCADRNGGCSH 614	
Db	542 GDYVYWTDWQRSSIERVHKRSAREVIIDQLPDLMLKATNVHRVIGSNPCABENGCGSH 601	
Qy	615 LCFTPHATRCGCGIGLELLSDMKTCIVPEAFIVFTSRAAHRISLETNNNDVAIPITGV 674	
Db	602 LCLYRPOGLRCACPIGPFELISDMKTCIVPEAFLLFSRRADIRISLETNNNNVAIPITGV 661	
Qy	675 KEASALDPDVSNHHIYWTDVSLKTIISRAFNMGSSVEHVVEGLDYPEGMAVDWVGKLYW 734	
Db	662 KEASALDPDVTDNRIYWTDISLKTIISRAFNGSALHVEHVVEGLDYPEGMAVDWVGKLYW 721	
Qy	735 ADTGTRNIEVARLDQGRQVILVWRDLNPRSLALDPTKGYIYWTWEGGKPRIVAFMDGT 794	
Db	722 ADTGTRNIEVSKLDQGRQVILVWKDLSPRALALDPAEGFYWTEWGGKPKIDRAAMDGS 781	
Qy	795 NCMTLVKVGANDLTIDYADQRLYWTDLTNWILESSNMLQOERVVADDLPHFPGLTQY 854	
Db	782 ERTTLVPNVGVRANGTLTIDYAKRRLYWTDLTNLIESSNMLGLNREVADDLPHFPGLTQY 841	
Qy	855 SDIYIYWTDNWLNHSTRADKTSGRNRTLIOGHLDFVMDILVFHSSROGLNDCMNNNGCG 914	
Db	842 QDIYIYWTDNWSSRIERANKTSGRNRTLIOGHLDFVMDILVFHSSRQSGWNECASNGHCS 901	
Qy	915 QLCILAIP-GGHRCCASHYTLTDPSSRNCSPPTTFLFSQKSAISRMIPDQHSPLDIPL 973	
Db	902 HCLAVPVGGFCVCPAHYSLNADNRCTCSA PTTTFLFSQKSAINRWVIDEQSQSDIILPI 961	
Qy	974 HGLRNVAIDYDPLDKFIYWDGQRN-IKRAKDGTQGF-VLTSLSQGNPDPRPHDLIS 1031	
Db	962 HSLRNVAIDYDPLDKQLYWTDNRQNMIRKAQEGSGQFTVVVSSVSPQNLEIQPYDLIS 1021	

Qy	1032 DIYSRTLFTWCEATNTINVHRLSGEAMGVVLGRDRDKPRATVVAERGYLYTNNQDRAA 1091	
Db	1022 DIYSRYIVTCEATNVINVTFLDGRSVGVVLKGQDRPRATVVAERGYLYTNNQDERSP 1081	
Qy	1092 KIERAALDGTERTEREVLFTGLIRPVALVVDNTLGLKFWVDADLKRIESCDSLGSANRLTLED 1151	
Db	1082 KIERAALDGTERTEREVLFFSGLSKPTALADSLRGLKFWADSLRRIESDSLGSANRIVLED 1141	
Qy	1152 ANIVQPLGLTLGLKHLWIDRQOQMIERVEKTDGKTRIOGRVAHLNGIHAVEVSELEE 1211	
Db	1142 SNLQPLVGLTVFENWLFWIDKQOQMIERIKDWTGREGRTKVOARLAQLSDIHAVKELNLOE 1201	
Qy	1212 FSAHPCARDNGGCSHICIAKGDGTTPRCSPVHLVLLQNLTLTCGEPPTCSPQFACATGEI 1271	
Db	1202 YRQHPCADNGGCSHICLVKGDGTTRCSCPMHLVLLQDELSCGEPPTCSPQOFTCTGEI 1261	
Qy	1272 DCIPGAWRCDFPECDQSDQDEBEGCPVCSAAQPPCARGOCVDLRCLRCDEACQDRSDEAD 1331	
Db	1262 DCIPVAWRCDFTECEDHSDDELNCPCVSESQFCASGQCIDGALRCNGDANCQDKSDEKN 1321	
Qy	1332 CDAICLPNPRCAGCQCVLKKQCCSPDCIDGSDLMCEITKPPSDDSPAHSAAIGPVI 1391	
Db	1322 CEVLCLLDQFRANGQCIGKHKCDHNVDCSKDELDCYTERP---APQATNTVGSVI 1378	
Qy	1392 GIILSLFVMGCVYFVQCQRYAGANGPPPEHY-VSGTTPHVLNFTIAPGSGHQGPFTG 1450	
Db	1379 GVIVTIFVSGTVYFICQMLCPKMGDGTMTNDYVHGVPASVPLGYVPHPSLGSPLPG 1438	
Qy	1451 IACKSMSSVSLMGRGGVPLYDRNHVTGASSSSSSSTKATLYPPLNPPPPSPATDPSL 1510	
Db	1439 MSRGKMISSLSIMGSSGPP-YDRAHVTGASSSSSSSTKGYFPAILNPPPPSPATERSH 1497	
Qy	1511 YNMDFYSSNIPATAR--PYRPYIIRGMAPPPTPCSTDVDCSDYSASR-----WKASKYV 1563	
Db	1498 YTEFGYSSNSPSYRHSYRPSYRHPAPPTTPCSTDVDCSDYAPSRMTSVATAKYT 1557	
Qy	1564 LDLNSDSDPPPPPTPHSQVLSAE---DSCPPSPATERSY-FHLFPFPPSPCTDSS 1615	
Db	1558 SDLNYSDEPVPPTTPRSQVLSAENYSCPPSPYTERSYSHHLYPPPPSPCTDSS 1613	

RESULT 6

LRP6 MOUSE

ID_LRP6_MOUSE STANDARD; PRT; 1613 AA.

AC O89572;

DT 05-JUL-2004 (Rel. 44, Created)

DT 05-JUL-2004 (Rel. 44, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Low-density lipoprotein receptor-related protein 6 precursor.

GN Name=Lrp6;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

ON NCBI_TaxID=10090;

RX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/c; TISSUE=Liver;

RX MEDLINE=98369644; PubMed=9704021; DOI=10.1006/bbrc.1998.9061;

RA Brown S.D., Twells R.C., Hey P.J., Cox R.D., Levy E.R., Soderman A.R.,

RA Metzker M.L., Caskey C.T., Todd J.A., Hess J.F.,

RT "Isolation and characterization of LRP6, a novel member of the low

RT density lipoprotein receptor gene family."

RL Biochem. Biophys. Res. Commun. 248:879-888(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.G., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heideh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

FT	DOMAIN	150	192	LDL-receptor class B 3.
FT	DOMAIN	194	235	LDL-receptor class B 4.
FT	DOMAIN	236	276	LDL-receptor class B 5.
FT	DOMAIN	282	324	EGF-like 1.
FT	DOMAIN	372	413	LDL-receptor class B 6.
FT	DOMAIN	415	456	LDL-receptor class B 7.
FT	DOMAIN	458	500	LDL-receptor class B 8.
FT	DOMAIN	502	540	LDL-receptor class B 9.
FT	DOMAIN	543	586	LDL-receptor class B 10.
FT	DOMAIN	588	628	EGF-like 2.
FT	DOMAIN	674	715	LDL-receptor class B 11.
FT	DOMAIN	717	758	LDL-receptor class B 12.
FT	DOMAIN	760	801	LDL-receptor class B 13.
FT	DOMAIN	803	840	LDL-receptor class B 14.
FT	DOMAIN	843	884	LDL-receptor class B 15.
FT	DOMAIN	889	930	EGF-like 3.
FT	DOMAIN	977	1024	LDL-receptor class B 16.
FT	DOMAIN	1026	1067	LDL-receptor class B 17.
FT	DOMAIN	1069	1112	LDL-receptor class B 18.
FT	DOMAIN	1114	1155	LDL-receptor class B 19.
FT	DOMAIN	1157	1197	LDL-receptor class B 20.
FT	DOMAIN	1203	1244	EGF-like 4.
FT	DOMAIN	1248	1286	LDL-receptor class A 1.
FT	DOMAIN	1287	1323	LDL-receptor class A 2.
FT	DOMAIN	1325	1361	LDL-receptor class A 3.
FT	DOMAIN	1459	1475	Poly-Ser.
FT	DOMAIN	1566	1573	Poly-Pro.
FT	DOMAIN	1603	1608	Poly-Pro.
FT	SITE	1487	1491	PPSP motif.
FT	SITE	1604	1608	PPSP motif.
FT	DISULFID	286	297	By similarity.
FT	DISULFID	293	308	By similarity.
FT	DISULFID	310	323	By similarity.
FT	DISULFID	582	603	By similarity.
FT	DISULFID	599	612	By similarity.
FT	DISULFID	614	627	By similarity.
FT	DISULFID	893	904	By similarity.
FT	DISULFID	900	914	By similarity.
FT	DISULFID	916	929	By similarity.
FT	DISULFID	1207	1218	By similarity.
FT	DISULFID	1214	1228	By similarity.
FT	DISULFID	1230	1243	By similarity.
FT	DISULFID	1249	1263	By similarity.
FT	DISULFID	1256	1276	By similarity.
FT	DISULFID	1270	1285	By similarity.
FT	DISULFID	1288	1300	By similarity.
FT	DISULFID	1295	1313	By similarity.
FT	DISULFID	1307	1322	By similarity.
FT	DISULFID	1326	1338	By similarity.
FT	DISULFID	1333	1351	By similarity.
FT	DISULFID	1345	1360	By similarity.
FT	CARBOHYD	42	42	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	81	81	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	281	281	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	433	433	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	486	486	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	692	692	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	859	859	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	865	865	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	926	926	N-linked (GlcNAc. .) (Potential).
FT	CONFLICT	83	83	S -> T (in Ref. 2).
FT	CONFLICT	317	317	M -> L (in Ref. 2).
FT	CONFLICT	586	586	V -> I (in Ref. 2).
FT	CONFLICT	622	622	G -> S (in Ref. 2).
FT	CONFLICT	933	933	S -> T (in Ref. 2).
FS	SEQUENCE	1613 AA;	180254 MW;	3C2ABC8EEB17622 CRC64;

Query Match 69.8%; Score 6097.5; DB 1; Length 1613;

Best Local Similarity 69.3%; Pred. No. 0;

Matches 1120; Conservative 216; Mismatches 259; Indels 21;

15 LILLALLCGCPAPAAASPLLLFANRDRVRLVDAGGVKLESTIVVSGLEDAAAVV

QY 15 LLLLLLALCGCPAPAAAAPLLLFAANRRDVRVLVDAGGVKLESTIVWSGLEDAAAVDFQFSK 74

Db 4 VLRSLACFCVLLRAA-PLLYANRRDLRLVDAATNGKENATVIVGGLEDAADAAVDFVGH 62
QY 75 GAVYWDVSEBAIKOTYLNQTAAGVQNVVIGLVSFDPDGLACDWGKLYWTDSETRIEV 134
Db 63 GLIYWSDVSEBAIKETENKS-ESVQNVVSGLLSPDGLACDWLGEKLYWTDSETRIEV 121
QY 135 ANLNGTSKRVLPWQDLOPRAIALDPAHYMYWTDWGETPRIERAGMDGSTRKIIVDSI 194
Db 122 SNLDGSLKRVLPWQDLOPRAIALDPSGFMWTDWGEVVKIERAGMDGSSRFVIINTEI 181
QY 195 YWPNGLTLDLEOKLYWADAKLSPFTHRANLDGSPKQVVEGSLTHPPALTISGDTLYWTD 254
Db 182 YWPNGLTLDYQERKLYWADAKLNFTHKSNLDGTRNQAUVVKSGLPFPALTIFEDTLYWTD 241
QY 255 WQTSIHACNKRKTGGRKEILSALYSPMDIOVLQSOERQFPFHTTRCEEONGGCSHLCLLSP 314
Db 242 WNTSHILACNKTGEGLEIHSNTPSPMDIHAFSQORPNATNPGCINDGGCSHLCLMSP 301
QY 315 SEPFTYCAPTGVQODNGRTCKAGAERVLRLARTDLRRISLDTPDFTDIVLQVDDIRH 374
Db 302 VKPFYQACPTGVKLMENGKTKCDGATELLRLARTDLRRISLDTPDFTDIVLQLEDIRH 361
QY 375 AIAIDYDPLEGVVWTDDEVAIRRAYLDGSAQTLVTEINPDGIAVDVARNLYWTD 434
Db 362 AIAIDYDPEVGYIYWTDDEVAIRRSFIDGSGSQFVVTQAHPDGIADVARNLYWTD 421
QY 435 TGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVMGLMYWTDGCKNPKICANLDGOER 494
Db 422 TGTDRIEVTRNGTWKILISEDLDEPRAIVLDPWGYMYWTDWGEIPKIERAALDGSDR 481
QY 495 RVLVNASLWPNGLALDLOEGKLYWGDARTKIEVINVDGTRKRTLLLEDKLPHPFGFTLL 554
Db 482 VVLVNTSLGWPNGLALDVEGTIYWGDAKTKIEVMNTDGTGRRVLVEDKIPHPFGFTLL 541
QY 555 GDFIYWTDWQRRSERVHKVVASRDVIIDQLPDLMLKANVAVKVVGNVPCADNRGGCSH 614
Db 542 GDIYVWTDWQRRSERVHKVSAEREVIIDQLPDLMLKATSVHRVIGSNPCAEONGGCSH 601
QY 615 LCFTEHATRCGPIGLLELSDMTKCIVPEAPLFTVPSAAIHRISLENNNDVAIPLTVG 674
Db 602 LCLYRQGLRCACPIGFLIGDMKTCIVPEAPLFSRRADIRRSLENNNNVAIPLTVG 661
QY 675 KEASALDPDVNNHIYWTDVSLKTSIRAFMNGSSVEHVVEFGLDYPEGMAVDMWGNKLYW 734
Db 662 KEASALDPDVNNRIYWTDVSLKTSIRAFMNGSALEHVVEFGLDYPEGMAVDMWGNKLYW 721
QY 735 ADTGTRIEVARLDGQFQVILVWRDLNPNRSLALDPTKGYIYTWEGGKPRIVAFMDGT 794
Db 722 ADTGTRIEVARLDGQFQVILVWRDLNPNRSLALDPAEGFMWTEWGGKPIDRAAMDGS 781
QY 795 NCMTLVKVGANDLTIDYADORLYWTDLTDTNMISSNMLQERVVIADDLPHFPGLTOY 854
Db 782 ERTTLVNVVGRANGLTIDYAKRULYWTDLTDTNLISSDMLGAINREVIADDLPHFPGLTOY 841
QY 855 SDIYIYWDNLHLSIERADKTSGRNRTLQGLHDFVMDILVFHSSRQGLNDGMHNGCG 914
Db 842 QDIYIYWDNWSRSIERANKTSQNRNTIIQGLHLDYMDILVFHSSRQAGWNECASNGHCS 901
QY 915 QLCIAIP-GGHCRCASHYLDPSRNCSPPTTLLFSQKSAISRMIPDDQHSPLILPL 973
Db 902 HLCIAVPVGGFVCGPAHYSLNADNRKTSAPSTELLFSQKSAINRWVIDEQSPDIILPI 961
QY 974 HGLRNVKAIYDPLDKFYWVWGRQN-IKRAKDDGTQPP-VLTSLSQGNPDPRPHLISI 1031
Db 962 HSLRNVKAIYDPLDKFYWVWGRQN-IRKAHEPDGGGFNVVANSVANQNLIEQPYLSI 1021
QY 1032 DIYSRTLPWTEATNTINVRHLSGEAMGVLRGRDKPRAIVNAERGLYFTVMQDRAA 1091
Db 1022 DIYSRYIYWTCEATNTIDVTRLDERSVGVLKGGQDRPRAIVNPEKGYMTFTNLQERSP 1081
QY 1092 KIERAALDGTREVLFTGLIRPVVALVVDNTLGLKFWVDADLKRIESCDLSGANRLTIED 1151
Db 1082 KIERAALDGTREVLFTGLIRPVVALVVDNTLGLKFWVDADLKRIESCDLSGANRLTIED 1141

QY 1152 ANIVQPLGLTILGKHLWYIDRQQOMIERVEKTTGDKRTRIOGRVAHLGTIHAVEVSLEE 1211
Db 1142 SNILQPLGLTIFENWLYWIDRQQOMIEKIDMTGREGRTKVQARIAQLSDIHAVKLNLOE 1201
QY 1212 FSAHPCARDNGGCSHICIAKGDGTPRCVHLVLLQNLATCGEPPTCSPDQFACATGEI 1271
Db 1202 YRQHPCAQDNGGCSHICLVKGDGTRCSCPMHLVLLQDELSCGEPPTCSPQOFTCTGTGI 1261
QY 1272 DCIFCANRCDPFPCDDOSDEGCPVCSAAOPPCARGOCVLDRLRLRCDEADQCQDSDEAD 1331
Db 1262 DCIFVANRCDGFTCEDHSDBLNCFVCSSESQFQCASGQCIDGALRCNGDANQCQDSDEKN 1321
QY 1332 CDATCLNPQFRCASQCQVLIKQCCDSFPDCLDGSDELCEITKPPSDPSAHSSAIGPVI 1391
Db 1322 CEVLCLDLOFCANGQCVKHKCDHSDVDCSDRDELDCYFTEEP---APQATNVGSI 1378
QY 1392 GIILSLFVMGVYFVQVQVQVYAGANGPPHEHYVSGTP-HVPLNFIAPGSGORHPFTG 1450
Db 1379 GVIITFVSGTIYFICQRMLCPRMKDGETWTDVNVVHSPASVPLGYVPHPSLSGSLPG 1438
QY 1451 IACGKSMSSVSLMGRGCVPLYDRNHVTGASSSSSSSTKATLYPPIILNPPSPATDPSL 1510
Db 1439 MSRGKSMISSLSINGSGSGPP-YDRAHVGTGASSSSSSSTKGTYPPIILNPPSPATERSH 1497
QY 1511 YNMDMFYSSNIPATAR--PYRPIYIRGMWAPPTPCSTDVCDSDYSASR-----WKASKY 1563
Db 1498 YTMFEGYSSNSPSTHRSYRPSYRHPAPPTPCSTDVCDSDVAPSRMTSVATAKGYT 1557
QY 1564 LDLNSDPPYPPPTPHSQYLSAE---DSCPPSPATERSY-FHLPPPPSPCTDSS 1615
Db 1558 SDVNYDEPVPPTPRSQYLSAENYESCPSPTERSYSHLYPPPPSPCTDSS 1613

RESULT 7

Q8AYFO PRELIMINARY; PRT; 1613 AA.
AC Q8AYFO; AC Q8AYFO;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Lipoprotein receptor-related protein 6.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
NC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22194715; PubMed=12204281; DOI=10.1016/S0925-4773(02)00205-8;
RA Houston D.W., Wylie C.;
RT "Cloning and expression of Xenopus Lrp5 and Lrp6 genes."
RL Mech. Dev. 117:337-342(2002).
DR EMBL; AF508961; AAN09807.1; -.
DR HSSP; P01130; 1AJJ.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002172; LDL receptor A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00005; EGF; 4.
DR Pfam; PF00057; Ldl_recept_a; 3.
DR Pfam; PF00058; Ldl_recept_b; 19.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00181; EGF; 4.
DR SMART; SM00192; LDLa; 3.
DR SMART; SM00135; LY; 20.
DR PROSITE; PS01209; LDLRA_1; 3.
DR PROSITE; PS00068; LDLRA_2; 3.
DR Lipoprotein; Receptor.
KW SEQUENCE 1613 AA; 180602 MW; BD93F09273DDDL77 CRC64;

Query Match	69.6%; Score 6084; DB 2; Length 1613;
Best Local Similarity	68.7%; Pred. No. 0;
Matches 1109; Conservative 206; Mismatches 280; Indels 20; Gaps 11;	
QY	16 LLLALLCGCPAPAAAPLLLPANRRDVRVLVADGVKLESTIVVSGLEDAADVDFQFSKG 75
DB	4 LLRSLLMCSLCLRASPLLLYANRRDLRVDTAGMKGNSIVVSGLEDAADVDFVSRG 63
QY	76 AVYWTDSBEAAIKOTYLNQTAQVNVVIGSLVSPDGLACDWMGKGLYWTDSNRIEVA 135
DB	64 LIYWSVDSEAAIKRDNKTGSS-QDVVIGLVSPDGLACDWMGKGLYWTDSNRIEVS 122
QY	136 NLNGTSRKVLFWQDLQDPRATLPAHGYMYWTWGETPRIERAGMGDSIRKIIIVDSDIY 195
DB	123 NLGSLRKVIFWQDLQDPRATLPAHGYMYWTWGETPRIERAGMGDSIRKIIIVDSDIY 182
QY	196 WPNGLTLDLEQKLYWADAKLSFTHIRANLDGSRFQKVVEGSLTHPPFALTLSGDTLYWTDW 255
DB	183 WPNGLTLDYDEQKLYWADAKLSFTHIRANLDGSRFQKVVEGSLTHPPFALTLSGDTLYWTDW 242
QY	256 QTRSIHACNKRGTGKRKEILSALYSYPMIDQVLSQERQFFHTRCEEDNGGCSHLCLLSPS 315
DB	243 TTRSIHACNKRGTGKRKEILSALYSYPMIDQVLSQERQFFHTRCEEDNGGCSHLCLLSPM 302
QY	316 EPPYTCACPTGVOLQDNGRTCKAGAEVLLILARETDLRRISLDTPTDVIQVDDIRHA 375
DB	303 EPPYTCACPTGVOLQDNGRTCKAGAEVLLILARETDLRRISLDTPTDVIQVDDIRHT 362
QY	376 IADYDPLEGVYVWTDDEVAIRAYLDGSAQTLVNTINDPGIADVWVARNLYWTDI 435
DB	363 IADYDPEGVYVWTDDEVAIRAYLDGSAQTLVNTINDPGIADVWVARNLYWTDI 422
QY	436 GTDRIEVRNLGTSRKILVSEDLDEPRALAHVYVGLMYWTDGNGPKIECANLDGQERR 495
DB	423 GTDRIEVRNLGTSRKILVSEDLDEPRALAHVYVGLMYWTDGNGPKIECANLDGQSDRI 482
QY	496 VLVNASIGWPNGLALDLOEGKLYWGDADTKIEVINVDGTERLLEDKLPHEFTLLG 555
DB	483 ILVNTSLGWPNGLALDLOEGKLYWGDADTKIEVINVDGTERLLEDKLPHEFTLLG 542
QY	556 DFIYWTWQRRSIRERHVKHAKASRDVIIDQLPDLMLKAVNVAKVVTGNTPCADRNGGCSHL 615
DB	543 DYVVTWQRRSIRERHVKHAKASRDVIIDQLPDLMLKAVNVAKVVTGNTPCADRNGGCSHL 602
QY	616 CFFTHPRCCGPTGLELLDMKTCIVPEAFVFTSRAAIIHRISELTNNNDVAIPLTGK 675
DB	603 CLYRPOGPRCACPIGLELLDMKTCIVPEAFVFTSRAAIIHRISELTNNNDVAIPLTGK 662
QY	676 EASALDFVSNHIIYWTDSVLSKTIISRAFMGSSVEHVVERGLDYPEGMAVDMGKLYWA 735
DB	663 EASALDFVSNHIIYWTDSVLSKTIISRAFMGSSVEHVVERGLDYPEGMAVDMGKLYWA 722
QY	736 DTGNTNREVARLDGQFVLRDLNDRSLALDPTKGIYVWTEWGGKPRIVRAFMDCGN 795
DB	723 DTGNTNREVARLDGQFVLRDLNDRSLALDPTKGIYVWTEWGGKPRIVRAFMDCGN 782
QY	796 CMTLVKVRANDLTIDYADQRLYWTDLDTNMISSNNMLGQERVVVADLPHPPGLTOYS 855
DB	783 RITLVDPVGRANGLTIDYAEERLYWTDLTDLNMISSNNMLGQERVVVADLPHPPGLTOY 842
QY	856 DIIYVTDWNLHSIRADKTSGRNRTLIQHLDFVMDILVHSSRODGLNDCWHNNGCGQ 915
DB	843 DIIYVTDWNLHSIRADKTSGRNRTLIQHLDFVMDILVHSSRODGLNDCWHNNGCHSH 902
QY	916 LCLALIP-GGHRGCGCASHYTLDPFSRNCSPPTTFLIFSOKSAISRMIPDDQHSPLILPLH 974
DB	903 LCLALIPSGYTCGCPVHSLTNDKTCAPSFLSFKNAINRMVLDGQSPDILPLIH 962
QY	975 GLRNVKADYDPLKFIYVWQGRONIKRAKDDGTQPF-VLTSLSQGNPDQPHDLSDI 1033
DB	963 NLRNVRAIDYDPLEKLYWIDSRONIRRAQEDSGSMTIVASTIPNQNMQPYDLSIDI 1022
QY	1034 YSRNLFWTCATNTINVHLSGEAMGVVLRGDRDKPRAIVVNAERGILYFTNMQDRAKI 1093

DB	1023 YSRNLFWTCATNTINVHLSGEAMGVVLRGDRDKPRAIVVNAERGILYFTNMQDRAKI 1082
QY	1094 ERAALDGTREVLFTTGLIRPVALLVNDNTLGLFWVDADLKRIBSCDLSGANRUTLEDAN 1153
DB	1083 ERAALDGTREVLFTTGLIRPVALLVNDNTLGLFWVDADLKRIBSCDLSGANRUTLEDAN 1142
QY	1154 IVQPLGLTILGKHLIYWDROOQMIERVEKTTGDKRTIQRVAHLTGIHAEVSVLEBFS 1213
DB	1143 VLPQVGLTFGNLYIWDROOQMIERIEKTRGVGRKTRIOARIPLLTDIHAVTELIMDEYK 1202
QY	1214 AHCARDNGGCSHICIAKGDTGTPRCPCPVHLVLIQ-NLLTCGEPTTSCPDQFACATGID 1272
DB	1203 EHPCSQNGGCSHICIVKGDGTRCSCPLHLVLLQDELSCEGPTTSCPLHFTCTGEVD 1262
QY	1273 CIPGAMRCDFPECDDBEGCPVCSAAPPCARGQCVDLRLRCDEADQCDSDEADC 1332
DB	1263 CIPGAMRCDFPECDDBEGCPVCSAAPPCARGQCVDLRLRCDEADQCDSDEKNC 1322
QY	1333 DAICLPNQFRCASQCVLIQKQCDSPDCIDGSDDELCEITKPPSDSDSPAHSAGPVIG 1392
DB	1323 NEVCAPDQFHCGSGQCIGKRRCDLSPDCSDSDDEQACYPTEEP---PPSTNTIGSIIG 1379
QY	1393 IILSLFVWGGVYVFCQVVCORVYAGANGPPPEY-VSGTPEHVPINFTIAPGSGHGPFTGI 1451
DB	1380 VILTFVVGGVYVFCQVLCRRMKGDGETWINDVYVHGVSVPLAYVPHPSLSLGLPDM 1439
QY	1452 ACCKSMSSVSLMGRCGVPLYDRNHVTGASSSSSSSTKATLYPILNPPSPATDPSLY 1511
DB	1440 SRGKSVIGSLIMAGSGGP-YDRAHVTGASSSSSSSTKATLYPILNPPSPATERSHY 1498
QY	1512 NMDMFYSNTPATAR--PYRPIIRGMAPPPTTCTDVCSDYS-----ASRWKASKYL 1564
DB	1499 TMEFGYSNTPSTRSYRPNYRHFAPPTTCTDVCSDYS-----ASRWKASKYL 1558
QY	1565 DLNDSDPYPPPTPHSQYLSAE---DSCPPSPATERSY-FHLPPPPSPCTDSS 1615
DB	1559 DLNDSBPVPPPTPHSQYLSAEENYSCPPSPATERSYSHHLYPPPPSPCTDSS 1613
RESULT 8	
Q95V09	PRELIMINARY; PRT; 1678 AA.
ID	Q95V09
AC	Q95V09
DT	01-DEC-2001 (TrEMBLrel. 19, Created)
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	Arrow.
GN	Name=arr;
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.
OX	NCBI_TaxID=7227;
EN	[1]
RP	SEQUENCE FROM N.A.
RA	Wehrli M., Dougan S.T., Caldwell K., O'Keefe L., Schwartz S.,
RA	Vaizel-Ohayon D., Schejter E., Tomlinson A., DiNardo S.;
RT	"arrow encodes an Ldl receptor-related protein essential for Wingless
RT	signaling in Drosophila.";
RL	Nature 0:0-0(2000)
DR	EMBL; AY005815; AAF91072.1; --
DR	HSSP; P01130; IN7D.
DR	FlyBase; FBgn000119; arr.
DR	GO; GO:0016055; P:Wnt receptor signaling pathway; IMP.
DR	InterPro; IPR000152; Asx hydroxyl_s.
DR	InterPro; IPR006209; EGF-like.
DR	InterPro; IPR006210; IEGF.
DR	InterPro; IPR002172; Ldl_receptor_A.
DR	InterPro; IPR000033; Ldl_receptor_rep.
DR	Pfam; PF00008; EGF; 2.
DR	Pfam; PF00057; Ldl_recept_a; 3.
DR	Pfam; PF00058; Ldl_recept_b; 16.

DR PRINTS; PRO0261; LDLRECEPTOR.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00192; LDLA; 3.
DR SMART; SM00135; LY; 20.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS01209; LDLRA_1; 2.
DR PROSITE; PS00068; LDLRA_2; 1.
DR PROSITE; PS00068; LDLRA_2; 1.
SQ SEQUENCE 1678 AA; 187438 MW; 4426A8CD1806BA18 CRC64;

Query Match 40.4%; Score 3527.5; DB 2; Length 1678;
Best Local Similarity 42.5%; Pred. No. 1.2e-212;
Matches 720; Conservative 289; Mismatches 542; Indels 143; Gaps 43;

QY 7 GPPWPLLLLLLL-----ALCGCPAPAAA-----SPLLLFANRDV 43
DB 39 GVGRQMLIGFLLCFGISNQYKVMPPSSSLIASP-PASAFVNTPATLFTTHDI 97
QY 44 RLVD-----AGVKLESTIVVSGLEDAADVPQFGKAVYVTDVSEEAIKQTYLQGT--- 96
DB 98 QVANITRPTGGPQID--VIVRDLAEMAIDPYAKNLVCMVTDGREIIECAQTSSALQP 155
QY 97 ---AAQNVNISGLVSDGLACDWGKLYWTDSETRIEVANLNGTSRKVLFQDLDOP 153
DB 156 LLRAPKQTVISTGLDKPEGLADHYTDKIYWDGKRIEIVATLDGRYQKLFWTDLDOP 215
QY 154 RAIALDPAHGYMYWTDGETPRIERAGMDGS--TRKIIIVSDIYWPNGLTIDLEEOKLYW 211
DB 216 RAVAVVAPKLLIWDNGEYFKIERASMDGDLRMTLVKEHVPWNGLAVDLKNELIYW 275
QY 212 ADAKLSFIHRANDGSPQKRVVEGSLTHPALTLSGDTLYWTDQWTRSIHACNKRGTGKR 271
DB 276 TDGKHFIIDVNRLDGSSRRITV--NNLKYPFSLTFYDDRLYWTDQWGLNALDIQT--REL 333
QY 272 KEILSALYSPMDIQVLSQEQPPFHTCEBDNGSCSHLCILSPSEPFYTCAPGTGVOLOD 331
DB 334 KELIDTPKAPNSVRAWDPSLQYEDNPFCHNNGNCSHLCILATNSKGFSCAPGTGVL--I 392
QY 332 NGRTCKAGAEVLLAARTRDLRLSLDTPDFTVLQVDDIRHAIAIDYDPLEGYVYWD 391
DB 393 SANTCANGSQEMPIVQRTQISLISLSDPYTIFPLPKGVKVAIADYDPEVEHIYWD 452
QY 392 DEVRAIRAYLDGSGAQTIVNTEINDPGIADVARNLYWTDGTDRIEVTRNGTSRK 451
DB 453 VETVTKRAHAGTGVTDFVTSEVRHPDGLDALNRLYWTDVTVDRIEVRCLDGTARK 512
QY 452 ILVSEDLDEPRAIALHPVWGLMYWTDGE--NPKTECANLDQERRVLNWSLHPNGAL 510
DB 513 VLIYHELEPRAIAVAPSLGWMFWSWNERKPKYERASLDGSERVLVSENLGPNGL 572
QY 511 DLQEGKLYWGDAKTDKIEVINVDGTRKRTLLDKLPHIFGFTLLAGDTFYTDMORRSIER 570
DB 573 DIEAKIYWCQKTDKIEVANMGSGRRVVISDNLKHFLGLSLDDLYWTDWRRSIDR 632
QY 571 VHKYKA--SRDVIIDQLPDLMLKAVNAVYVGTNTPCADRNGCGCSHLCTFPHATRCGCP 629
DB 633 AHKITGNRIYVVDQYFDLMGLKVTRLREVGQACAVRNGCGCSHLCLNRPDYVCRCAI 692
QY 630 GLELLSDMKTCIVPEAFIVFTSRAAIIHISL---ETNNNDVAIPLTGVKEASALDFVSN 686
DB 693 DYELANDKRTCVVPAALFLFSRQEHISIEYNEGHNHDERIPFKDVRDAHALDVSAE 752
QY 687 NHIYVTVSLKTIISRAFMNGSVHVEFGDYDPEGMADVMMGNLYWADTGTNRIVAR 746
DB 753 RRIYVTVQKSKCIFRALNGSYVQRIYVDSGLIGDGIADVLANNIYSDAEARIEVAR 812
QY 747 LDGQFQVLRDLNPRSLALDPTKGIYIYWTGKGPRIYVRAFMDGTNCMTLVKDYKGRA 806
DB 813 LDGSSRRVLLWKGVESRSLVLEPRRGVYWTB--SPTDSIRRAAMDGSLQTIIVAGANHA 871
QY 807 NDLTIVADQRLYV--TDLTDNMISSENNLQOER--VVIADLLPHFPFGTQYSDIYIYWDN 864
DB 872 AGLTFDQETRLRYWATQSRPAKIESADWQKRIQLVGSMDDEPYAVSLQDYVYVWSDN 931

QY 865 LHSIERADKTSGRNRTLIQGHLDVFMVMDILVFHSSRQDGLDCMHNNGCGCOLCIAIPG-- 922
DB 932 TGDIERVHKTTGQNRSLVHSGMTYITSLVFNDRKQTVGNPCKVYNGGCSHLCLAQPGRR 991
QY 923 GHRGCGASHYTLDPSSRNCSPPTTFLFSPQSAISRMIIPDQHSPLDILPLHGRUNKAI 982
DB 992 GMTCACTHYQLAKDGVSCIPPRNYIIFSORNSGRLLPNTTDCPNPLPVSQ--KNIRAV 1050
QY 983 DYDPLDKFIYWDGR--ONIKRAKDDGTQPFVLTLSLQSQNDPDROPHDLSDIYSRTLPWT 1041
DB 1051 DYDPTTHIIWIEGRSHIRKSLANGTKVSLAN-----SGQFFDLAIDIIGRLLEFWT 1103
QY 1042 CEATNTINVRHLSGEAMGVVLRGDRKPRAIIVNAERGIYLYTNMQDRAAKIERAALDGT 1101
DB 1104 CSQNSINVTSLFSGSVGIDTGDSEKPRNIAVHAMKRLLEFWTDVGSQA--IIRAVDGN 1162
QY 1102 EREVLTFTGLIRPVVALVVDNTLGLFWVDADLKRIBSCDLSGANRLTLEDANIVQPLGLT 1161
DB 1163 ER--VELAYKLEGVLTALDQSDMIY--AHGKRIDAIDINGKNKTLVSMHISQVINIA 1219
QY 1162 ILGKHLWIDRQOQMIERVEKTTGDKETRIQGRVAHLTGTHAVEEVSLEEFSAHPCARDN 1221
DB 1220 ALGQFVYWLQ--DKTGVBERI--TVNGERSAELOQLUPQITDIRAVWTPDPKVLNRNTHCMHSR 1277
QY 1222 GGCSHICIAKXGDTGR-----CSCPVLHLVLLQNLLTCGEPPTCSPDQFACATG-----E 1270
DB 1278 TKCSHICIASGEGIARTFVCSCKHMLLEDKENCGAFACGPDHFTCAAPVSGISDVN 1337
QY 1271 IDCIPTGAWRCGPECDQDDEBGCVPVCSAAQPPCARGCQVDLRLRCDGEADCDQDRDEA 1330
DB 1338 KDCIPASWRCGQKDCPKSDEVGCTCRADQFSCQSGECIDKSLVCDGTTNCANGHDEA 1397
QY 1331 DCAIAC--LPNQFCASGQ--CVLIIKQCCSDPDCCI DGSDE---LMCEITKPPSDSDSPAHS 1385
DB 1398 DC---CKRGEFQCPINKLCISAALLCDGWENCADGADSSDICLQRRMAPATKRAFMI 1454
QY 1386 AIGPVIIGIILSLFVGGVYFVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 1445
DB 1455 LIGATMITIFS-----VYLL--QFCRTRIGKSRTEPKDQDQATDPLSPSTL-----SKS 1501
QY 1446 GPPTGIACGSMMSV--SLMGRGGVPLVDNRNVTGASSSSSTKATLYPPILNPPPS 1503
DB 1502 QRVSQKIA---SVADAVRMSTLNSRNSMNSYDRNHTGASSSTTNGSSMVAYP--INPPPS 1556
QY 1504 PATDPSLYNMDMFYSSNI PATAREPYRVIIRGMAPPTTTCSTDYC---DSDSA----- 1554
DB 1557 PAT-----RSRRPYRHYKINQPPPTTCTDIDCDESDSNYTSKSNNN 1600
QY 1555 SRWKASY-----YLDLNSDSDFYPPPTPHSQYLS-----AEDSCPPSPATERS-YF 1601
DB 1601 SNGGATKSSSSAAACLOQYGVSEYPPPTPRSHYHSDVRIVPESCCPPSPSSRSSTYF 1660
QY 1602 HLPFPSPSPCTDSS 1615
DB 1661 SPLPPSPSPVQSPS 1674

RESULT 9
Q9V6Q0 PRELIMINARY; PRT; 1678 AA.
AC Q9V6Q0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE CG5912-PA.
GN Name=arr; ORFNames=CG5912;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.


```
QY 571 VHKYKA-SRDVIIDQLPDLMLGLKAVNVAKVVGTTNCPCADRNGCGCSHLCPFFTPHATRCGCP 629
DB 633 AHKITGNRIIVVDDYQYDPLMGLKVTLRLEVRGQACAVRNGCGCSHLCLNRPDRVVCRAI 692
QY 630 GLELLSMKTCIIVPEAFVLVFTSRAAIIHISL---ETNNNDVAIPLTGVKEASALDFVSN 686
DB 693 DYELANDKRTCVVPAALFLFSRQEHIGRISTEYNEGNHNDERIPFKVDRAHALDVSAE 752
QY 687 NHIYTVDSLTKTIKRAFNGSSVHVVEFLDYPEGMADVMMGNLKYADWTGTRIEVAR 746
DB 753 RRIYTDQSKCIFRAFLNGSYVORIYVDSGLIGPDGIAVDWLANNIYSDAEARIEVAR 812
QY 747 LDGQFQVLVWRLDNPRLSALDPTKGIYIYTWEGGPRIVRAFMGDTNCMTLVLDKVGRA 806
DB 813 LDGSSRRVLLWKGEPRSLVLEPRRGYMYWTE-SPTDSIRRAAMDGSLQTIIVAGANHA 871
QY 807 NDLTIDVADQRLYV-TLDTNMISSNMLGQER-VVIADDLPHFPGLTQYSDIYIYTDWN 864
DB 872 AGLTFDQETRLYVATQSRPAKIESADWDGKKQILVGSMDPEYAVSLYQDYVYVYSDWN 931
QY 865 LHSTERADKTSGRNRTLIQGHLDPMVDILVHSSRQDGLNDGMNNGOCGOLCLAIPIG-- 922
DB 932 TGLIERVHKTQGNRSLVHSGMTYITSLVFNDRQRTQVNPCKVNNGCGCSHLCLAQGPERR 991
QY 923 GHRGCGASHYTLDPSSRNCSPPTFLFSQKSAISRMIPDDQHSFDLILPLHLGLRNKAI 982
DB 992 GMTACPTHYQLAKDGVSCIPPRNYIIFSQRCFGRLLPNTTDCNPILPVSIG-KNIRAV 1050
QY 983 DYDPLKFIYVVDGR-QNIKRAKDDGTQFPVLTSLSQGNPDPRDHLSDIYSRTLFWT 1041
DB 1051 DYDPIITHIYIEGRSHSIRKSLANGTKVSLAN-----SQQFPDLAIDIIGRLFWT 1103
QY 1042 CEATNTINVRHLSCEANGVLRGDRDKPRAIVVNAERGVLFTYTNMQDRAAKIERAALDGT 1101
DB 1104 CSQSNINVTSLFSGVIGDITGSEKPRNIAVHAMKRLLLFWTDVGSQA-IIRARVDGN 1162
QY 1102 EREVLFTTGIRPVALVVDNTLGLKFWVDADLKRIESCDSGANRLTLEDANIVQPIGLT 1161
DB 1163 ER-VELAYKLEGVTAALDOQSDMIY--AHGKRIDDAIDINGKKNKTLVSHHSQVINIA 1219
QY 1162 ILGKHLTWIDRQQOMIERVKTQDKTRIQGRVAHUTGTHAVEEVSLEBSAHPCARDN 1221
DB 1220 ALGSGFYVWLD-DKTGVERI-TVNGERSAELQRLPQITDIRAVVTPDPKVLNRHNTCMHSR 1277
QY 1222 GGSCHICIAKGDGPR---CSCPVLHVLQNLITCGEPPTCSPDQACATG-----E 1270
DB 1278 TKSHICIASGEGARTDVCSCPKHMLLEDKENCGAFACGPDHFTCAAPVSGISDVN 1337
QY 1271 IDCIPGAWRCDFPECDQSDDEGCPVCSAAQFPFCARGQCQVDLRLCDGADCDRSDSEA 1330
DB 1338 KDCIPASWRCDGQKDCPKSDEVGCPCTRADQFSCQSGECIDKSLVCDGTTNCANGHDEA 1397
QY 1331 DCAIIC-LPNQFRCASQ-CVLIKQCCDSPDCIDGSDR---LMCEITKPPSDSDSPAHSS 1385
DB 1398 DC---CKRPGFQCPINKLCITSAALLCDGWENCADGADDESSDICILORMAPATDKRAFI 1454
QY 1386 AIGPVIGIILSLFWGCVYVFCQVRVQRYAGANGPEPHEVVSCTPHVPLNFIAPGCSQH 1445
DB 1455 LIGATMITIFI-----VILL---QFCRTRIGSRTEPKDQATDPLSPSTL-----SKS 1501
QY 1446 GPFTGIACGSMMSV--SLMGGRGVPLVDRNHVTGASSSSSTKATLYPPTLNPPPS 1503
DB 1502 QRVSKIA--SVADAVRMSTLNSNSWNSYDRNHITGASSSTGSSWVAIP--INPPPS 1556
QY 1504 PATDPSLYNMDMFYSSINATARYPIYIRGMAPPPTTPCSTDVC---DSDYSA----- 1554
DB 1557 PAT-----RSRRPYHYKIINQPPPTTPCSTDICDSDSDNYTSKSN 1600
QY 1555 SRWKASKY-----YLDLNSDSPPYPPPTPHSQYLS-----AEDSCPPSPATERS-YF 1601
DB 1601 SNGGATKSHSSAAACLOQGYDSEPPYPPPTPRSHYHSDRIVPESSCPFPSPSRSTYF 1660
QY 1602 HLFPFPPSPCTDSS 1615
```

```
DB 1661 SPLPPFSPVQSPS 1674
RESULT 10
Q9NH9 PRELIMINARY; PRT; 1678 AA.
AC Q9NH9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE LDL-related protein LR66.
GN Name=arr;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Avery B.J., Tsang G., Evans-Holm M., Skarnes W.C., Rubin G.M.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF223365; AAF28358.1; -.
DR HSP; P01130; IN7D.
DR FlyBase; FBgn000119; arr.
DR GO; GO:0016055; P:Wnt receptor signaling pathway; IMP.
DR InterPro; IPR00152; Asx_hydroxyl_S.
DR InterPro; IPR00209; EGF_like.
DR InterPro; IPR00210; IEGF.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR00033; Ldl_receptor_rep.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00057; Ldl_recept_a; 3.
DR Pfam; PF00058; Ldl_recept_b; 16.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00192; LDLA; 3.
DR SMART; SM00135; LY; 20.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS01209; LDLRA_1; 2.
DR PROSITE; PSS0068; LDLRA_2; 1.
SQ SEQUENCE 1678 AA; 187483 MW; C8891D89E7F7F3AA CRC64;

Query Match 40.3%; Score 3521.5; DB 2; Length 1678;
Best Local Similarity 42.4%; Pred. No. 2.9e-212;
Matches 719; Conservative 288; Mismatches 544; Indels 143; Gaps 43;

QY 7 GPPWPLLLLLLLLL-----ALCGCPAPAAA-----SPLLLFANRDRV 43
DB 39 GVGWRQMLIGFLLCFGISNSWQYKNVHMPSSSLIASP-PASAFVNTPATLLFTTRHDI 97
QY 44 RLVD-----AGVKLESTIVVSGLEDAADVPQFSGAVYTDVSEBAIKQTYLNQTG--- 96
DB 98 QVANITRPTGGPQID--VIVRDLAEANMAIDFYAKNLVCMWTDSGREIECAQTSSALQP 155
QY 97 ---AAVQNVVIGLVPDGLACDVGKLYWTDSETNRIEVANLNGTSRKVLFWDLDLP 153
DB 156 LLRAPKTVISTGLDKEGLAMDWTOKIYWDGKRIEVAITLDGRYQKVLFWTDLDP 215
QY 154 RAIALDPAHGYMYTWDGETPRIERAGMOGS--TRKIIVDSDIYWPNGLTIDLEEQLYW 211
DB 216 RAVAVVPARKLLIITWDGVEPKIERASMDGDLPLSRMTLVKEHVPWNGLAVDLNELIYW 275
QY 212 ADAKLSIHRANLDGSPQKVEGSLTHPALTUSGDTLTWTDWQTSIHACNKRGTGKR 271
DB 276 TDGKHHPIDVWLDGSSRRITV--NNKYFPFSLTFYDDRLYWTDWQSGSLNALDLQT-REL 333
QY 272 KEILSALYSMDIQVLSQERQPPHTRCEBDNGSCSHLCLLSPSEPFVTCACPTGVQLQD 331
DB 334 KELIDTPKANSVRAMDPISLPQYEDNPNCAHNGNCSHLCLLATNSQGSFCAPTGVKL-I 392
QY 332 NGRTCKAGAEVLLARRTDLRRISLDPDTIVLQVDDDIRHAIAIDYDPLEGVVYWD 391
```

393 SANTCANGQEMFIVQRTQISILSDSPDYTFIPPLGKVKYAIADYDPVEEHIYNSD 452
392 DEVRAIRAYLDGSAQTLNTEINDDPGIAVDWVARNLYWTDGTDRIEVTRLNGTSRK 451
453 VETYIKRAHADGTGTVTFVSEVRHPDGLADMLARNLYWTDVIXDRIEVCRLDGTARK 512
452 ILVSDLEDEPRALALHPVYGLMYWTDWGE-NPKIECANLDGOERRVLYNASIGWENGLAL 510
513 VLIYEHLEPRALAVAPSGLMFWSDMNERPKVERASLDGSRVVLVSENGLWPNGLAL 572
511 DLQEGKLYWGDAKTDKIEVINVDGTRKRTLLKLEPHIFGFTLLGDFTYWDWQRRSIER 570
573 DIEAKALYWCDDKTKIEVANWMDGSGRRVWISDNLKHLFGLSILDYLYWTDWQERSIDR 632
571 VHKVKA-SRDVLIIDQLPDLMLKAVNAKVVTNTPCADRNGGCSHLCPFTTHATRCGCP 629
633 AHKITGNRRIVVVDQYPLMGLKVTRLREVRQNAACAVRNGGCSHLCLNRPDRYVCRCAI 692
630 GLELLSDMKTICVPEAFIVFTSRAAIHRISL---ETNNNDVAIPLTGVKEASALDPOVSN 686
693 DYELANDKRTCVVPAFLFRQEHIGRISIEYNEGNHDERIPPKDVRDAHALDVVAE 752
687 NHIYWDVSLTKTISRAFNMGSSVEHVFEGLDYPEGMAVDMWGNLYWADTGNRIEVAR 746
753 RRIYWDQKSKICIPRAFLNGSVQRIVDSGLIGPDGIAVDMLANNIYWSDAEARIEVAR 812
747 LDGQFQVLRWDLNPRSLALDPTKGIHYTWGCKPRIIVRAFMGNGNCTMLDKVQGRA 806
813 LDGSRRLVLLMKGVEPRSLVLEPRGYMYTE-SPTDSIRRAAMDGDQLTIVAGANHA 871
807 NDLTIDYADQLYV-TDLDTNMISNNMLOGER-VVIADDLPHPPGLTQYSDYLYWTDWN 864
872 AGLTFDQTRLYWATQSRPAKIESADWDGKRQLVGSMDDEPVAVSLYQDYVWSDWN 931
865 LHSIERADKTSGRNRTLQGHLDVMDILVPHSSRQDGLNDCMHNNGCGOLCLAIPO-- 922
932 TGDIERVHKTQGNRLSVHSGWTVYITSLVFNDRKQGTGVNPKVNGNGCSHLCLAQPCR 991
923 GHRCCASHYTLDPSSRNCSPPTFLFSQKSAISRMTPDQHSPLDILPLHGLRNVKAI 982
992 GMTACAPHYOLAKDGVSCIPFRNYIIFSQNCRGLLPNTDPCNPLPVSQ-KNIRAV 1050
983 DYDPLDKFIYVDGR-QNIKRAKDDGTQFVFLTSLSQONPDROPDHLSDIYSRTLPWT 1041
1051 DYDPTTHIYIEGRSHSIEKSLANGTVKSLAN-----SGQPFDLAIDIIIGLLFPWT 1103
1042 CEATNTINVHLSGEMGVVLRGDRDKPRAIVNAERGVLVFTNMQDRAAKIERAALDGT 1101
1104 CSQSNINVTSLFSGVGVITDGSSEKPRNIAVHAMKRLLEFWTDVGSQA-IIRARVDGN 1162
1102 BREVLFITGLIRPVALVVDNTLGLFWADLKRIESCDLSGANRLTLEDANIVQPLGLT 1161
1163 ER-VELAYKLGVTALADQQSDMIY--AHGKRIDAIDINGKNKTLVSHHISQVINIA 1219
1162 ILGKHLVWIDRQQQMIERVETKTRTRIQRVAHLTGHAIVEVSLEEPSAHPCARDN 1221
1220 ALGSGFVWLD-DKTGVERI-TVNGERSAEQLRPLQITDIRAVWTPDPKVLNHTCHSR 1277
1222 GCSHICIAKGGTPR-----CSCPVLVLLQNLTTCGEPPTCSPPQFACATG-----E 1270
1278 TKSHICIASGEGARTDRVDCPCPKHMLLDEKENCAGFAACGPDHFTCAAPVSGISDVN 1337
1271 IDCIPGAWRCGPFECDDSDGECPCVCSAAQFPCAGQCVDLRLCDGEADCODRDEA 1330
1338 KDCIPASWRCDQKCDPKSDDEVGCTCRADQFSCQSGECIDKSLVCDGTTNCANGHDEA 1397
1331 DCDAIC-LPNQFRCASGO-CVLIKQCCDSFPDIDGSDSDE---LMCEITKPPSDSDPAHSS 1385
1398 DC---CKRPGFQCPINKLCISAALLCDGWENCADGADDESSDICLQRMARPATDKRAFI 1454
1386 AIGPVGIIILSLFWGVIYVQVQVQRYAGANGPPPHVVSQTPHVPFLNFIAPGGSQH 1445
1455 LIGATMITIFS-----VYLL---QFCRTRIGKSRTEPKDDQADTDLPSPTL-----SKS 1501

1446 GPFTGIACGKSMMSV--SLMGGRGVPLYDRNHVHTGASSSSSSTKATLYPPILNPPPS 1503
1502 QRVSKEIA---SVADAVRMSTLNSRSMNSYDRNHITGASSSTTNGSSMVAYP--INPPPS 1556
1504 PATDPSLYNMDMFYSSNIPATAPRVRYIIRGMAPPTTPCSTDYV---DSYSA----- 1554
1557 PAT-----RSPRYRHYKIINQPPPTPCSTDICDESDSNYTSKSNNN 1600
1555 SRWKASKY-----YLDLNSDSPPPPPPPHSQYLS-----AEDSCPPSPATERS-YF 1601
1601 SNGGATKXSSSSAAACLYGYDSEPPPPPTPRSHYHSDVRIVPSSCPCPPSPSSSTYF 1660
1602 HLFPPPPSPCTDSS 1615
1661 SPLPPPPSPVQSPS 1674

RESULT 11

Q7FV65 PRELIMINARY; PRT; 1698 AA.
AC Q7FV65;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000020132 (Fragment).
GN Name=ENSANGG00000017643;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA01008986; EAA00402.2; -.
DR HSSP; Q07954; 1D2L.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000152; Aax_hydroxyl_S.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; LDL_receptor_rep.
DR InterPro; IPR001005; Myb_DNA_binding.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00057; Ldl_recept_a; 3.
DR Pfam; PF00058; Ldl_recept_b; 17.
DR PRINTS; PR00261; LDLRECEPTOR.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS01209; LDLRA_1; 2.
DR PROSITE; PS00068; LDLRA_2; 1.
DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
FT NON TER.
SQ SEQUENCE 1698 AA; 187081 MW; 2498484BD5044E93 CRC64;

Query Match 37.3%; Score 3262.5; DB 2; Length 1698;
Best Local Similarity 40.7%; Pred. No. 5.8e-196;
Matches 696; Conservative 277; Mismatches 580; Indels 157; Gaps 44;

Qy 27 APAASP---LLLFANRPDRVLV-----DAGGVKLESTIVSGLEAAAVDFQSGAV 77
Db 21 ATAASTSTILLFTTYVDLRVLNVSVQPDNTALTYMDVLQLQDLNEASAMDFHARGLV 80
Qy 78 YWTDVSEALIKQTYNLQNTGAQVQ--NVVISGLVSPDGLACDWGKKLYTWDSETNRIEVA 135
Db 81 CWTENTLEVTCQCTNGTGLSVTSKTTVTITDGLDKPEGLAIDWYTDKLYWTDGESNRIEVA 140
Qy 136 NL-----NGTSRKVLFWODLQDPAIALDPAHGYMYWTDWG 171
Db 141 QLEPFGGGGAGATGGGGGIGGSGKQPNARLQKVLINSDLPQRAIALVPARRYMIWTDWG 200


```
QY 273 EILSALYSPMDIQVLQSOEROPFFHTRCEEDNGGCSHLCLLSSEBFFYTCACPTGVQLQDN 332
Db IIRNKLHFPMDIHTLHPQRPAGKRCGDNNGGCTHCLCL--PSGQNYTCACPTGFR-KIN 732
QY 333 GRTCKAGAEVLLARTRDIRISLTDPTDFTDVLQVDDIRHAIAIDVDPLRGYVYVWDD 392
Db 733 SHACAQLSKFLLFARRMDIRISFDTEDLSDVPLADVRSAVALDMSRDDHVIWTDV 792
QY 393 EVRAIRRAYLDGSGAQTLVNTEINDPGIADVWVARNLYWTDGTDRIEVTRNGTSRKI 452
Db 793 STDTISRAKWDGTGQKVVVDTSLESAGLAIDWVTKNLYWTDAGTRIEVANTDGSMTV 852
QY 453 LVSEDLDEPRATLHPVWGLMYWTDGKPNKTECANLQGERVLVNASLGNPGLALDL 512
Db 853 LIWENLDRPRDIVBPMGQYMYWTDGASPKTIERAGMDASNRQVITISNLTWPNGLAID 912
QY 513 QEGKLYWDGAKTDKLEVINVDGKTRTLEDKLPHIFGFTLLGDFIYWDQRRRIERVH 572
Db 913 GSORLYWADAGMKTIEFAGLDGSKKVLIGSQLPHFGTLTGQRIYWTDWQTSIQSAD 972
QY 573 KVKA-SRVDIIDLPLMGLKAVNVAKVGVNTPCADRNGGCSHLCTFFTHAT--RCGCP 629
Db 973 RLTLGLDRETLOENLENLMDIHVFHRQRPVITPCAVENGCGCSHLCLRSPSPGSCCTCPT 1032
QY 630 GLELLSDMKTICV-PEAFIVETSRAAIHRI SLETN-NDVAIPLT-GYKEASALDFVSN 686
Db 1033 GINLLDGTICPGMNSFLIFARRIDRVMSLDIPIYFADVVVPIWMTKNTIAIGVDPLE 1092
QY 687 NHIYWDVSLKTIISRAFNGSSVEHVFGDYPGMAVDMWGNLYWADGTNRIEVAR 746
Db 1093 GKVYWSDSLTHIRISASLDGSHEDIITTLGTLTDLAVDAIGRKVYWTDGTNRIEVGN 1152
QY 747 LDCQFQVLVWDLNPRSLALDPTKGIYIYTWGCKPRIIVRAFMDNGNCTLV-DKVR 805
Db 1153 LDGSMKRLVWQNLSPRAIYLHEMGFYMTDGENAKLERSGMDGSDRTVLINNLGW 1212
QY 806 ANDLTDYADQRYWTDLTWNIESNNLMGQERVVIADLPHPLGLTQYSDYIYWTWNL 865
Db 1213 PNLTVDTKSQLWADHAETRIEVLADGANRHTLVSPQHPGLTLLSDYIYWTDWQT 1272
QY 866 HSIERADTKSGNRNRTIQGLHDLFVMDILVFFHSSROGLNDQWNNNGCQGLCLAIPIGHR 925
Db 1273 RSIHRADKSTGNSVTLVRSNLPGLMDIOAVDRAQPLGKNGSRNGGCSHLCLRPSGFS 1332
QY 926 CGCASHYTLDPSSRNCSP-PTFLLFSSQSAISRMIPDDQHSPLILPLHGLRNKAIDY 984
Db 1333 CACPTGIQLKGDGKTCDSPEYLLFSSRGSIRISLTDTHDTHVHPVPGLNVISLDY 1392
QY 985 DPLDKFIYWDGRON-IKRAXDDGTQPPVLTSLSQGNPDQPHDLSIDIYRSLFWTCE 1043
Db 1393 DSDVGKYYTDFLDVIRADLNGSN--METVIGHGL---KITDGLADVWVARNLYWTD 1447
QY 1044 ATNTINVRHLSGEAMGVVLRGDRDKPRAIVVNAERGYLEFTNMQDRAAKIERAALDGT 1103
Db 1448 GRNTIEASRLDGSCKRLINNSLDEPRAIYVPRKGLFWTDW-GHIAKIERANLDGSR 1506
QY 1104 EVLFTTGLIRVALVWNTLKLFWADADLKRIESCDLSGANRLTLEDANIVQPLGLTIL 1163
Db 1507 KVLINADLGNPGLTLDYDTRIIYWVDADLHRIEADLNGKLQVLY-SHVSHPFALTOQ 1565
QY 1164 GKHLWIDRQQOMIERVEKTTGDKRTRIQRVAHLTGTHAVEEVSLE-EPSAHPCARDNG 1222
Db 1566 DRWIYWTDWQTSQRVDKYSGRNKETV--LANVEGLMDIIVSPQRQTGNAAGVNG 1622
QY 1223 GCSHICIAKGDTGRCSFVHLVLLQNLTLTCGPPTCSPDQFACATGEIDCIPG 1276
Db 1623 GCSHLCFARASDFV-CACP-----DEPDSDHPCS-----LVPG 1653
```

RESULT 14

Q76LU2

PRELIMINARY;

PRT; 1905 AA.

ID Q76LU2

```
AC Q76LU2;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE LDL receptor-related protein 13.
GN Name=LRP13;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Tian Q., Suzuki T., Okano A., Usuda N.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB073317; BAD18061.1; -.
DR HSSP; P01130; 1AJJ.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR000152; Asx_Hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00008; EGF; 4.
DR Pfam; PF07645; EGF_CA; 1.
DR Pfam; PF00057; Ldl_recept_a; 8.
DR Pfam; PF00058; Ldl_recept_b; 19.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00181; EGF; 7.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00192; LDL; 8.
DR SMART; SM00135; LY; 20.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; UNKNOWN 1.
DR PROSITE; PS01209; LDLRA_1; 7.
DR PROSITE; PS00068; LDLRA_2; 8.
DR EGF-like domain; Receptor.
KW EGF-like domain; Receptor.
SQ SEQUENCE 1905 AA; 211928 MW; 4A9ACB7D00ED54FE CRC64;

Query Match 30.7%; Score 2687; DB 2; Length 1905;
Best Local Similarity 43.0%; Pred. No. 1.1e-159;
Matches 539; Conservative 222; Mismatches 443; Indels 50; Gaps 21;

QY 33 PLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQFSKGAVYVWTDVSEEAIKOTYL 92
Db 440 PVLLFANRIDIRQVLPH--RSEYTLNLLNLENAIALDFHRRRLVFWSDVTLDRILRANL 497
QY 93 NOTGAAVQNVVISGLVSPDGLACDWGKKLYWTDSETNRLEVANLNGTSRKVLFWQLDQ 152
Db 498 N--GSNVEEVVSTGLESFGLAVDWVHDKLYWTDSGTSRIEVANLGDGHRKVLWQSLK 555
QY 153 PRAIALDPAGVYVWTDWGTPTRIERAGMDGSTRKIIVDSDIYWPNGLTIDLEBQKLYWA 212
Db 556 PRAIALHPMGTYIWTWNTPTRIEASMDGSGRIIADTHLFWPGLTIDYAGRIWV 615
QY 213 DAKLSFTHIRANLDGSRQKVVESGLTHPFAITLSGDTLYWTDWQTSRIHACNKTGGKRK 272
Db 616 DAKHHVIERANLDGSHRKAVISQGLPHFPFAITVFEDSLYWTDWHTKSINSANKFTGNQE 675
QY 273 EILSALYSPMDIQVLQSOEROPFFHTRCEEDNGGCSHLCLLSSEBFFYTCACPTGVQLQDN 332
Db 676 IIRNKLHFPMDIHTLHPQRPAGKRCGDNNGGCTHCLCL--PSGQNYTCACPTGFR-KIN 732
QY 333 GRTCKAGAEVLLARTRDIRISLTDPTDFTDVLQVDDIRHAIAIDVDPLRGYVYVWDD 392
Db 733 SHACAQLSKFLLFARRMDIRISFDTEDLSDVPLADVRSAVALDMSRDDHVIWTDV 792
QY 393 EVRAIRRAYLDGSGAQTLVNTEINDPGIADVWVARNLYWTDGTDRIEVTRNGTSRKI 452
Db 793 STDTISRAKWDGTGQKVVVDTSLESAGLAIDWVTKNLYWTDAGTRIEVANTDGSMTV 852
```

Qy	453	LVSEDLDEPRALHPVWGLMYTWDGKNGPKIECANLDGQERRVYNASLQWPNGLDLD	512	RP	CONCEPTUAL TRANSLATION.
Db	853	LJWENLDRPRDIVVEPGWYWTWDGASPKIERAGMDASNRQVLISSNLTPWGLAIDY	912	RL	Blatter M.-C.;
Qy	513	QEGKLYGDAKDKIEVINVDGTRKRTLLLEKLPHEIFGFTLLGDFIYWTWQORSIERVH	572	RN	Unpublished observations (MAR-2004).
Db	913	GSORLYWADAGMKITFAGLDGSKRVKLVIGSOLPHFPGLTLYGQRIYWTWQTSIQASD	972	RP	[2]
Qy	573	KVKA-SRDVIIDQIPDLMLKANVAVKVGTPCADRNGGCSHLCFFTPHAT--RCGCP	629	RC	SEQUENCE OF 375-1950 FROM N.A.
Db	973	RLTGLDRETLQENLEMDIHVFRQRPVFTTPCAVENGGCSHLCRSPPSGFSCTCPT	1032	RC	TISSUE=Brain;
Qy	630	GLELLSDMKTCIV-PEAFVPTSRAAIHRISLETNN-NDVAIPLT-GVKEASALDDFVSN	686	RC	MEDLINE=98360089; PubMed=9693030; DOI=10.1006/geno.1998.5341;
Db	1033	GINLLDGKTCSPGNSFLIFARIDVWVSLDIPYFADVVVPINMTKNTIALGVPLE	1092	RA	Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
Qy	687	NHIVYTDVSLTKISRAFPMNGSSVHVVEFGLDYPEGMAVDMGKNLYWADGTNRIVAR	746	RA	"Identification of high-molecular-weight proteins with multiple EGF-
Db	1093	GKVWSDTLHRISRASLDGSHQEDIIITGLQTTDGLAIDAIGKVVYTDGTNRIEVGN	1152	RA	like motifs by motif-trap screening.";
Qy	747	LDGQFQVLRDLDNRSALDPTKGIYVTEGGRPIVRAFWDTGNTCMTLV-DKVR	805	RA	Genomics 51:27-34 (1998).
Db	1153	LDGSMRKVLVWQNLDSRAIYLVHEMGPMYWDGNAKLSRSGDSDRTVLINNLGW	1212	RA	[3]
Qy	806	ANDLTIDYADQRLYWTDLDTNMIESNNLGOERVVIADDLPHFPGLTQYSDIYIYWTWNL	865	RA	SEQUENCE OF 1764-1950 FROM N.A.
Db	1213	PNGLTVDKTSQLLWADAHTERIEVADLNGANRHTLVSPVQHPGLTLLDSYIYWTWQT	1272	RA	TISSUE=Muscle;
Qy	866	HSIERADKTSRNTLITQHLDFVMDILVFHSSRDQGLNDCHMNGGQGLCLAIPOGHR	925	RA	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
Db	1273	RSIHRADKSTGNNVILRSNLPGLMDIQAVDRAQPLGKNGSRGGCSHLCCLPSPGFS	1332	RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Qy	926	CGCASHYLDPSRNCSP-PTTFLLFSQSAISRMIPODQHSPLILPLHGLRNVKIDY	984	RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Db	1333	CACPTGLQKDGDKTCDSPETYLFSRGSIRISLTDHTDTHVFPVPGANNVISLDY	1392	RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Qy	985	DPLDKFYVWDGRON-IGRAKADGTQFPVLTSLSGQGNPDROPHDLSDIIVSRLEWTC	1043	RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Db	1393	DSVDGKYTYTDFVDIRADLNGSN--MEVIGHGL---KTTDGLAVDWVARNLYTDT	1447	RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Qy	1044	ATNTINVRHLSGEAMGVVLRGDRDKPRAIVVNAERGVLFTYFTNMQDRAAKIERAALDGT	1103	RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Db	1448	GRNTIEASRLDGSKVKVLINNSLDEPRAIAVFPKGLYFTDW-GHIAKIERANLDGSR	1506	RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Qy	1104	EVLFTTGLIRPVAVDNTLGLFWADLKRIESCDLSGANRLTLEDANIVQPLGLTIL	1163	RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Db	1507	KVLINADLGNWGLTLDYTRIIYVWDAHLDRIESADLNGLRQVLV-SHVSHPFALTQ	1565	RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Qy	1164	GKHLWIDRQOQMIERVEKTKGDKRTRIQGRVAHLTGHAVEEVSLE-EFSAHPCARDNG	1222	RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Db	1566	DRWIYTDWQTSQRVDKYSGRNKETV---LANVEGLMDIIVSPQRQTGTNACGVNG	1622	RA	Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Qy	1223	GCSHCIAKGDTPRCSCPVLHVLQNLTLTGCEPPTCSPDQAFACATGEIDCIPG	1276	RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Db	1623	GCSHLCFARASDFV-CACP-----DEPDSPHPCS-----LVPG	1653	RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RESULT 15				RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
LRP4 HUMAN				RA	Butterfield Y.S.N., Krywinski M.I., Skalska U., Smalls D.E.,
ID -LRP4 HUMAN				RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;
AC 075056;				RT	"Generation and initial analysis of more than 15,000 full-length human
DT 05-JUL-2004 (Rel. 44, Created)				RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
DT 05-JUL-2004 (Rel. 44, Last sequence update)				RL	-!- FUNCTION: Potential cell surface endocytic receptor, which binds
DE Low-density lipoprotein receptor-related protein 4 precursor (Multiple				CC	and internalizes extracellular ligands for degradation by
DE epidermal growth factor-like domains 7).				CC	lysosomes.
GN Name=LRP4; Synonyms=KIAA0816, MEGF7;				CC	-!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
OS Homo sapiens (Human).				CC	-!- TISSUE SPECIFICITY: Expressed in several regions of the brain.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				CC	-!- SIMILARITY: Belongs to the LDLR family.
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				CC	-!- SIMILARITY: Contains 3 EGF-like domains.
OX NCBI_TaxID=9606;				CC	-!- SIMILARITY: Contains 8 LDL-receptor class A domains.
[1]				CC	-!- SIMILARITY: Contains 19 LDL-receptor class B domains.
				CC	-!- CAUTION: The sequence has been constructed according to mouse and
				CC	rat sequences. The N-terminus part differs from mouse and rat
				CC	sequences, but is confirmed by EST.
				CC	-----
				CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
				CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
				CC	the European Bioinformatics Institute. There are no restrictions on its
				CC	use by non-profit institutions as long as its content is in no way
				CC	modified and this statement is not removed. Usage by and for commercial
				CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/
				CC	or send an email to license@isb-sib.ch).
				CC	-----
				DR	EMBL; AC021573; -- NOT ANNOTATED_CDS.
				DR	EMBL; AB011540; BAA32468.1; --
				DR	EMBL; BC031360; AAH37360.1; --
				DR	EMBL; BC041048; AAH41048.1; --
				DR	HSSP; P00736; 1AQP.
				DR	Genew; HGNC:6696; LRP4.
				DR	H-InvDB; HIX0009607; --
				DR	MIM; 604270; --
				DR	GO; GO:0005509; F:calcium ion binding; NAS.
				DR	InterPro; IPR000152; Asx hydroxyl S.
				DR	InterPro; IPR001881; EGF_Ca.
				DR	InterPro; IPR006209; EGF_like.
				DR	InterPro; IPR002172; LDL_receptor_A.
				DR	InterPro; IPR000033; Ldl_receptor_rep.
				DR	Pfam; PF00008; EGF; 5.
				DR	Pfam; PF00057; Ldl_recept_a; 1.
				DR	Pfam; PF00058; Ldl_recept_b; 19.

DR SMART; SM00179; EGF CA; 1.
 DR SMART; SM00135; LV; 20.
 DR PROSITE; PS00010; ASX HYDROXYL; 1.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 3.
 DR PROSITE; PS00026; EGF_3; FALSE_NEG.
 DR PROSITE; PS01187; EGF CA; 1.
 DR PROSITE; PS01209; LDLRA_1; 8.
 DR PROSITE; PS00068; LDLRA_2; 8.
 KW Calcium-binding; EGF-like domain; Endocytosis; Glycoprotein; Receptor;
 KW Repeat; Signal; Transmembrane.
 FT SIGNAL 1 20
 FT CHAIN 21 1950
 FT DOMAIN 21 1768
 FT TRANSMEM 1769 1791
 FT DOMAIN 1792 1950
 FT DOMAIN 71 112
 FT DOMAIN 115 151
 FT DOMAIN 154 189
 FT DOMAIN 192 228
 FT DOMAIN 235 271
 FT DOMAIN 275 311
 FT DOMAIN 314 350
 FT DOMAIN 356 395
 FT DOMAIN 399 439
 FT DOMAIN 440 479
 FT DOMAIN 525 566
 FT DOMAIN 568 609
 FT DOMAIN 611 653
 FT DOMAIN 655 696
 FT DOMAIN 698 738
 FT DOMAIN 743 782
 FT DOMAIN 830 871
 FT DOMAIN 873 914
 FT DOMAIN 916 958
 FT DOMAIN 960 1001
 FT DOMAIN 1002 1042
 FT DOMAIN 1138 1179
 FT DOMAIN 1181 1222
 FT DOMAIN 1224 1266
 FT DOMAIN 1268 1308
 FT DOMAIN 1309 1350
 FT DOMAIN 1442 1483
 FT DOMAIN 1485 1526
 FT DOMAIN 1528 1570
 FT DOMAIN 1572 1617
 FT SITE 1611 1614
 FT DISULFID 72 89
 FT DISULFID 79 102
 FT DISULFID 96 111
 FT DISULFID 116 128
 FT DISULFID 123 141
 FT DISULFID 135 150
 FT DISULFID 155 167
 FT DISULFID 162 180
 FT DISULFID 174 188
 FT DISULFID 193 205
 FT DISULFID 200 218
 FT DISULFID 212 227
 FT DISULFID 236 248
 FT DISULFID 243 261
 FT DISULFID 255 270
 FT DISULFID 276 288
 FT DISULFID 283 301
 FT DISULFID 295 310
 FT DISULFID 315 327
 FT DISULFID 322 340
 FT DISULFID 334 349
 FT DISULFID 357 369
 FT DISULFID 364 382
 FT DISULFID 376 394
 FT DISULFID 403 414

FT DISULFID 410 423
 FT DISULFID 425 438
 FT DISULFID 444 454
 FT DISULFID 450 463
 FT DISULFID 465 478
 FT DISULFID 747 758
 FT DISULFID 754 767
 FT DISULFID 769 781
 FT CARBOHYD 309 309
 FT CARBOHYD 543 543
 FT CARBOHYD 764 764
 FT CARBOHYD 946 946
 FT CARBOHYD 1122 1122
 FT CARBOHYD 1460 1460
 FT CARBOHYD 1512 1512
 FT CARBOHYD 1512 1512
 FT CONFLICT 1599 1599
 FT CONFLICT 1691 1691
 FT CONFLICT 1907 1907
 SQ SEQUENCE 1950 AA; 215963 MW; D298624D70B2A287 CRC64;
 Query Match 30.7%; Score 2686; DB 1; Length 1950;
 Best Local Similarity 42.8%; Pred. No. 1.4e-159;
 Matches 541; Conservative 229; Mismatches 447; Indels 46; Gaps 21;
 QY 33 PLLLFANRRDRLVDAGVGKLESTIVVSGLEDAADVPQSKGAVYWDVSEBAIKQTYL 92
 DB 485 PVLFFANRDIRQRLPH--RSEYTLNLENAIDLPFHRRELVPFMSDVTLDRLRANL 542
 QY 93 NOTGAAQNVVISGLVSPDGLACDWGKLYWTDSETRIEVANLNGTSRKVLWQDLDQ 152
 DB 543 N--GSNVEEVVSTGLESPGGLAVDWDVHDKLYWTDSTGSTRIEVANLDGAKRKLWQNLK 600
 QY 153 PRAIALDPAGHYWYWDWGTPTRIERAGMDGSTRKIIIVDSIYWPNGLTIDLEQKLYWA 212
 DB 601 PRAIALHPMEGTIYWDWGTPTRIEASSMDGSGRRIIADTHLPWNGLTIDYAGRRMYWV 660
 QY 213 DAKLSFTHRANLDGSRQKVEGSLTHPPFALTLSGDTLYWTDQTRTHACNKTGGKKR 272
 DB 661 DAGHVIERANLDGSHRKAVISQGLPHFPALTVEFDSLYWTDWHTKINSKNTGKNQE 720
 QY 273 EILSALYSPMDIQVLSQEROFFFTRCEEDNGGCSHLCLLSPSEFFYTCAPTQVQLQDN 332
 DB 721 IIRNKLHPMDIHTLHPQRPAGKRNKRGDNGGCTHCL--PSQNYTCAPTQGR-KIS 777
 QY 333 GRTCKAGAEVLLARITDLRLSLDTPDFTDIVLQVDIRHAIDAIDYDPLEGVYVWTD 392
 DB 778 SHACAOQLDKFLFARRMDIRISFDTEILSDDDVIPLADVRSVAVALDWDSDRDHYWTDV 837
 QY 393 EVRAIRRAYLDGGAOTLVNTEINDDPGIADVWVARNLYWTDGTGTRIEVTRNGTSRKI 452
 DB 838 STDTISRAKWDGTQGVVVDTSIESPAGLAIDWVTKLYWTDAGTDRIEVANTDGSRTV 897
 QY 453 LVSEDLDEPRAIALHPVMGLMYWTDNGENPKIECANLDGQERRVNLNAGSPNGLALDL 512
 DB 898 LIWENLDRPRDIIVPEPMGGYWYWDGASPKIERAGMDASGRQVLISSNLTPNGLAIDY 957
 QY 513 QEGKLYWGDAKTKIEVINVDGTRKRTLLDEKLPHPFGTLLGDFIYWTQWRRIERHV 572
 DB 958 GSORLYWADAGMKTIEPAGLDGSKRKVLIQSOLPHFPGLTYGERIYWTQWTKSIQASD 1017
 QY 573 KVKA-SRDVIIDQLPDLMLGKAVNVAKVCTNPCADRNGGCSHLCPFTPHAT--RCQCP 629
 DB 1018 RLTLGDRETIQENLENLMDHVHRRPPVSTPCAMENGCSHLCLSPNPSGSCCTPT 1077
 QY 630 GLELLSDMKTCTIV-PEAFVFTSRAAHRISLETNN-NDVAIFLT-GVKEASALDFDVSN 686
 DB 1078 GINLLSDGKTCSPQWNSFLIFARRIDIRMYSLDIPYFADVVVPINITMKNTIAIGVDPQE 1137
 QY 687 NHIYWTDVSLKTIISRAFMNGSSVHVVEFGLDYPEGMADVDMGNLYWATGTINRIEVAR 746
 DB 1138 GKVYWSDSLTHIRSRANLDGSHQEDIIITGLQTDDGLAVDAIGRKVYWTDTGTNRIEVGN 1197

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2005, 01:24:44 ; Search time 58 Seconds
(without alignments)
2679.139 Million cell updates/sec

Title: US-09-931-375A-2
Perfect score: 8740
Sequence: 1 MEAAPPGPPWPLLLLLLLLL.....TERSYPFLPPPPSPCTDSS 1615

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8736	100.0	1615	JR0372	low density lipopr
2	6153.5	70.4	1613	JR0272	low density lipopr
3	6097.5	69.8	1613	JR0273	low density lipopr
4	2654	30.4	527	JR0373	low density lipopr
5	2193	25.1	4544	S02392	alpha-2-macroglobu
6	2180	24.9	4543	A53102	alpha-2-macroglobu
7	2113	24.2	4545	S25111	alpha-2-macroglobu
8	1856	21.2	4660	T42737	gp330 protein prec
9	1458	16.7	4753	A47437	LDL-receptor-relat
10	1024	11.7	252	T46336	hypothetical prote
11	914.5	10.5	1847	T13308	probable vitellog
12	913.5	10.5	1984	T13171	probable vitellog
13	902	10.3	1650	S53457	dominant autoantig
14	824.5	9.4	1207	1 EGHU	epidermal growth f
15	805	9.2	1217	1 EGM5MG	epidermal growth f
16	794	9.1	2643	T29149	hypothetical prote
17	753	8.6	1133	1 EGRT	epidermal growth f
18	735	8.4	2215	2 T00348	LR11 protein - mou
19	663	7.6	2180	2 T29764	hypothetical prote
20	651.5	7.5	863	1 S51789	VLDL receptor prec
21	648.5	7.4	873	1 QRBRVD	VLDL receptor prec
22	648	7.4	873	1 A49729	VLDL receptor prec
23	643	7.4	873	1 I48952	VLDL receptor prec
24	634.5	7.3	1661	2 T31330	head-activator bin
25	617.5	7.1	837	1 A29512	LDL receptor prec
26	611.5	7.0	996	2 JR0237	apolipoprotein E r
27	608	7.0	909	1 QRXL11	LDL receptor 1 pre
28	607.5	7.0	862	1 QRMSLD	LDL receptor precu
29	606	6.9	860	1 QRHULD	LDL receptor precu

ALIGNMENTS

RESULT 1

JR0372 low density lipoprotein receptor related protein - human

C:Species: Homo sapiens (man)

C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004

C:Accession: JR0372; JE0329

R:Kim, D.; Inagaki, Y.; Suzuki, T.; Ioka, R.X.; Yoshioka, S.Z.; Magoori, K.; Kang, M.; Ch

J. Biochem. 124, 1072-1076, 1998

A:Title: A new low density lipoprotein receptor related protein, LRP5, is expressed in he

A:Reference number: JR0372; MUID:99054722; PMID:9832610

A:Accession: JR0372

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1615 <KIM>

A:Cross-references: UNIPROT:075197; DDBJ:AB017498

R:Dong, Y.; Lathrop, W.; Weaver, D.; Qiu, Q.; Cini, J.; Bertolini, D.; Chen, D.

Biochem. Biophys. Res. Commun. 251, 784-790, 1998

A:Title: Molecular cloning and characterization of LR3, a novel LDL receptor family prote

A:Reference number: JR0329; MUID:9908902; PMID:9790987

A:Accession: JR0329

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1329, 'A', 1331-1615 <DON>

A:Cross-references: GB:AF077820; NID:G3831747; PIDN:AACT2791.1; PID:G3831748

C:Superfamily: low density lipoprotein receptor-related protein; EGF homology; LDL recept

F:605-640/Domain: EGF homology <EGF>

F:1259-1295/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F:1298-1332/Domain: LDL receptor ligand-binding repeat homology <LDL2>

F:1336-1370/Domain: LDL receptor ligand-binding repeat homology <LDL3>

Query Match 100.0%; Score 8736; DB 2; Length 1615;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MEAAPPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLPANRRDRLVLDAGGVKLESTIVVS	60
Db	1	MEAAPPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLPANRRDRLVLDAGGVKLESTIVVS	60
Qy	61	GLEDAAAVDFQSKGAVYWTDVSEAIKQTYLNQTGAAVQNVVIGSLVSPGLACDWGK	120
Db	61	GLEDAAAVDFQSKGAVYWTDVSEAIKQTYLNQTGAAVQNVVIGSLVSPGLACDWGK	120
Qy	121	KLYWTDSETNRNRIEVANLNGTSRKVLFWQDLQDPAIALDPAHGYWYTDWGETPRIERAG	180
Db	121	KLYWTDSETNRNRIEVANLNGTSRKVLFWQDLQDPAIALDPAHGYWYTDWGETPRIERAG	180
Qy	181	MDGSTRKLIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPQKRVGSLTTHP	240
Db	181	MDGSTRKLIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPQKRVGSLTTHP	240
Qy	241	FALTSLSGDTLYTWDQWTRSHACNKRKTGGRKEILSALYSPMDIQVLQSRQPPFHTRCE	300
Db	241	FALTSLSGDTLYTWDQWTRSHACNKRKTGGRKEILSALYSPMDIQVLQSRQPPFHTRCE	300

Db 241 FALTLSGDTLYWTDNQTRSIHACNKRKTKGKKEIILSALYSMPMDIQVLSQERQPPFHTTCE 300
Qy 301 EDNGGCSHLCLLSPSEPFYTCACPTGVQLQDNNGRTCKAGAEVLLARRTDLRRISLDT 360
Db 301 EDNGGCSHLCLLSPSEPFYTCACPTGVQLQDNNGRTCKAGAEVLLARRTDLRRISLDT 360
Qy 361 DFTDVLQVDDIRHAIADYDPLEGYVYVWTDDEVRAIRRAYLDGSGAQTLVNTEINDPDG 420
Db 361 DFTDVLQVDDIRHAIADYDPLEGYVYVWTDDEVRAIRRAYLDGSGAQTLVNTEINDPDG 420
Qy 421 IAVDMVARNLYWTDGTDRIEVTRNGTSRKLTVSEDLDEPRAIALHPVGMGLMYWTDGGE 480
Db 421 IAVDMVARNLYWTDGTDRIEVTRNGTSRKLTVSEDLDEPRAIALHPVGMGLMYWTDGGE 480
Qy 481 NPKIECANLQGERRVLVNASLGNWGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL 540
Db 481 NPKIECANLQGERRVLVNASLGNWGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL 540
Qy 541 LEDKLPHIFGFTLLGDFTYWTDQRRSIEVHVKKVKSADVDIIDLQPLMLGLKAVNAVW 600
Db 541 LEDKLPHIFGFTLLGDFTYWTDQRRSIEVHVKKVKSADVDIIDLQPLMLGLKAVNAVW 600
Qy 601 GTNPCADRNNGCSHLCTFTPHATRCGCPGLELLSDMTKTCIVPRAFLVFTSRAAHRISL 660
Db 601 GTNPCADRNNGCSHLCTFTPHATRCGCPGLELLSDMTKTCIVPRAFLVFTSRAAHRISL 660
Qy 661 ETNNNDVAIPITGVKEASALPDFVSNHHIYWTDSLTKTSIRAFMNGSSVEHVFEGLDYP 720
Db 661 ETNNNDVAIPITGVKEASALPDFVSNHHIYWTDSLTKTSIRAFMNGSSVEHVFEGLDYP 720
Qy 721 EGMADVWNGKLYWADTGTNRIEVARLDGQFQVLRVLDLNPRLALDPTKGYIYWTW 780
Db 721 EGMADVWNGKLYWADTGTNRIEVARLDGQFQVLRVLDLNPRLALDPTKGYIYWTW 780
Qy 781 GCKPRIVAFNDGTNCMTLVKVGANDLTIDYADQRLYTDLDTNMIESNMLGQERVV 840
Db 781 GCKPRIVAFNDGTNCMTLVKVGANDLTIDYADQRLYTDLDTNMIESNMLGQERVV 840
Qy 841 IADDLPHFPGLTQYSDIYIYWTDNLHLSIERADKTSGRNRTLQGHLDPMVDILVFHSRQ 900
Db 841 IADDLPHFPGLTQYSDIYIYWTDNLHLSIERADKTSGRNRTLQGHLDPMVDILVFHSRQ 900
Qy 901 DGLNDCHNNGCGOLCLAIPEGHRCGCASHVTLDPSSRNCSPPTTLLFSQKSAISRMI 960
Db 901 DGLNDCHNNGCGOLCLAIPEGHRCGCASHVTLDPSSRNCSPPTTLLFSQKSAISRMI 960
Qy 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFYWVDGRQNIKRAKDDGTQPFVLTSLSGQ 1020
Db 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFYWVDGRQNIKRAKDDGTQPFVLTSLSGQ 1020
Qy 1021 NPDQPHDLSIDYSRTLFTWCEATNTINVRHLSGEAMGVVLGRDRKPRAIIVNAERG 1080
Db 1021 NPDQPHDLSIDYSRTLFTWCEATNTINVRHLSGEAMGVVLGRDRKPRAIIVNAERG 1080
Qy 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRISCD 1140
Db 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRISCD 1140
Qy 1141 LSGANRLTLEDANIVQPLGLTLGKHLWIDRQQQMIEVEKTTGDKRTRIQGRVAHLTG 1200
Db 1141 LSGANRLTLEDANIVQPLGLTLGKHLWIDRQQQMIEVEKTTGDKRTRIQGRVAHLTG 1200
Qy 1201 IHAVEVSLIEFSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLQNLLTCGEPTCS 1260
Db 1201 IHAVEVSLIEFSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLQNLLTCGEPTCS 1260
Qy 1261 PDQFACATGEIDCIPGAWRCDFGFECDQSDQSEEGCPVCSAAQFPCARQCQVDLRLCDDGE 1320
Db 1261 PDQFACATGEIDCIPGAWRCDFGFECDQSDQSEEGCPVCSAAQFPCARQCQVDLRLCDDGE 1320
Qy 1321 ADCQDRSEADCDAILCPNQFRCSGQCCLIKQCCDSFPDCIDGSDLMCBITKPPSDDS 1380
Db 1321 ADCQDRSEADCDAILCPNQFRCSGQCCLIKQCCDSFPDCIDGSDLMCBITKPPSDDS 1380

Qy 1381 PAHSSAIGPVGIIILSLFVMGVYVFCORVVCORYAGANGPFPHEYYSGTTPHVLNFIAP 1440
Db 1381 PAHSSAIGPVGIIILSLFVMGVYVFCORVVCORYAGANGPFPHEYYSGTTPHVLNFIAP 1440
Qy 1441 GGSQHGFPTGTIACGKSMSSVSLMGGRGVPLDYDRNHVGTGASSSSSSSTKATLPPILNP 1500
Db 1441 GGSQHGFPTGTIACGKSMSSVSLMGGRGVPLDYDRNHVGTGASSSSSSSTKATLPPILNP 1500
Qy 1501 PPSPATPDLNMDMFYSSNIPATAPRVPYIIRGMAPPPTPCSTDVCDSDYSASRWKAS 1560
Db 1501 PPSPATPDLNMDMFYSSNIPATAPRVPYIIRGMAPPPTPCSTDVCDSDYSASRWKAS 1560
Qy 1561 KYILDLNSDSDPYPPPTPHSQAISAEDESCPPSPATERSYFHLFPFPPSPCTDSS 1615
Db 1561 KYILDLNSDSDPYPPPTPHSQAISAEDESCPPSPATERSYFHLFPFPPSPCTDSS 1615

RESULT 2

JE0272

low density lipoprotein receptor-related protein 6 - human

C:Species: Homo sapiens (man)

C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004

C:Accession: JE0272

R:Brown, S.D.; Twells, R.C.J.; Hey, P.J.; Cox, R.D.; Levy, E.R.; Soderman, A.R.; Metzker,

Biochem. Biophys. Res. Commun. 248, 879-888, 1998

A:Title: Isolation and characterization of LRP6, a novel member of the low density lipoprotein

A:Reference number: JE0272; MUID:98369644; PMID:9704021

A:Accession: JE0272

A:Molecule type: mRNA

A:Residues: 1-1613 <BRO>

A:Cross-references: UNIPROT:O75581; GB:AF074264; NID:G3462526; PIDN:AAC33006.1; PID:G3462

C:Genetics:

A:Gene: LRP6

A:Map position: 12p11-12p13

C:Superfamily: low density lipoprotein receptor-related protein; EGF homology; LDL receptor

F:286-323/Domain: EGF homology <EGF1>

F:592-627/Domain: EGF homology <EGF>

F:1207-1243/Domain: EGF homology <EGF2>

F:1249-1285/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F:1288-1322/Domain: LDL receptor ligand-binding repeat homology <LDL>

F:1326-1360/Domain: LDL receptor ligand-binding repeat homology <LDL3>

Query Match 70.4%; Score 6153.5; DB 2; Length 1613;

Best Local Similarity 70.2%; Pred. No. 0;

Matches 1134; Conservative 205; Mismatches 256; Indels 21; Gaps 12;

Qy 15 LLLLLALCGCPAPAAAPLIPANRRDVRIVDAGGVKLESTIVVSGLEDAADVDFOSK 74

Db 4 VLRSLLACSFVLLRAA-PLLLVYANRRDLRLVDATNKENATIVVSGLEDAADVDFESH 62

Qy 75 GAVYTWDSBEAIKQTYLNOTGAQVNVVIGSLVSPDGLACDWVGKLYWTDSETNRIEV 134

Db 63 GLIYMSDVSEAIKRTFNKT-ESVQNVVVSGLLSPDGLACDWLGEKLYWTDSETNRIEV 121

Qy 135 ANLNGTSRKLFWQDLDPRAIALDPAHGYWYTWDTWGETPRIERAGDGGSTRKLIIVSDI 194

Db 122 SNLDGSLRKLFWQELDPRAIALDPSSGFMYTWDTWGEVFKIERAGDGGSTRFIINSEI 181

Qy 195 WYPNGLTIDLEOKLYWADAKLSFIHRANDLGGFPRQKRVGSLTHPFTALTLSGDTLYWTD 254

Db 182 WYPNGLTIDLEOKLYWADAKLSFIHRANDLGGFPRQKRVGSLTHPFTALTLSGDTLYWTD 241

Qy 255 WQTRSIHACNKRKTKGKKEIILSALYSMPMDIQVLSQERQPPFHTTCEEDNGCSHLCLLSP 314

Db 242 WSTHISILACNKYTGEGRLREIHSIDFSPMDIHAFTSQRPNATNPGIDNGCSHLCLMSP 301

Qy 315 SEPYTTCACGTGVLQDNNGRTCKAGAEVLLARRTDLRRISLDTPTDFTIVLQVDDIRH 374

Db 302 VKPFYQACPTGVKLENGTKCKDGAATELLARRTDLRRISLDTPTDFTIVLQVDDIRH 361

Qy 375 AIAIDYDPLEGYVYVWTDDEVRAIRRAYLDGSGAQTLVNTEINDPDGIAVDMVARNLYWTD 434

Db 362 AAIADYDPEVGIYWTDDVEAIRRSFIDGSGSQFVVTQAIAHPDGIADVWVARNLYWTD 421
Qy 435 TGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVMGLMYWTDWGENPKICANLDGOER 494
Db 422 TGTDRIEVTRNGTWKILISEDLDEPRAIVLDPMVGMYWTDWGEIPKIERAALDGSR 481
Qy 495 RVLVNASLWPNGLALDQEGKLYWGDAKTDKIEVINVDGTRKRTLLEDKLPHPFGFTLL 554
Db 482 VVLVNTSLGWPNGLALDYDEKLIYWGDAKTDKIEVMNTDGTGRRVLVEDKIPHPFGFTLL 541
Qy 555 GDFIYWTWQBRSTIERVHKVYKASRDVLIIDQLPDLMLKAVNVAKVGTNPNCDNRGCGSH 614
Db 542 GDIYVWTWQBRSTIERVHKVYKASRDVLIIDQLPDLMLKATVHVRVIGSNPCAEENGCGSH 601
Qy 615 LCFPTPHATRCGCPGLELLSDMTKICVPEAFVFTSRAAIHRISLETNNNDVAIPITGV 674
Db 602 LCLYRPOGLRCACPIGELISDMTKICVPEAFLEFRRADIRISLETNNNDVAIPITGV 661
Qy 675 KEASALDPDVSNHHIYWTDSVLSKTSIRAFMNGSSVHVVEFGLDYPEGMAVDMWGNKLYW 734
Db 662 KEASALDPDVTDNRRIYWTDSVLSKTSIRAFMNGSALHVVVEFGLDYPEGMAVDMWGNKLYW 721
Qy 735 ADTGTRIEVARLDQCFQVILWRDLNPRSLALDPTKGYIYWTWGWGKPRIVRAPMDGT 794
Db 722 ADTGTRIEVSKLDGQHRQVILWKLDSRALALDPAEGFMYTGWGKPKIDRAANDGS 781
Qy 795 NCMTLVKVGKANDLTIDYADORLYWTDLTDMNIESSNMLGOERVLTADLPHPGGLTOY 854
Db 782 ERTTLVNVVGRANGLTIDYAKRRLYWTDLDTNLISSNMLGNREVLTADDLPHPGGLTOY 841
Qy 855 SDIYVWTWNLHLSIERADKTSGRNRTIIOGHLDVMDILVPHSSRQDGLNDCMHNQCG 914
Db 842 QDIYVWTWMSRSIERANKTSQNRITIOGHLDVMDILVPHSSRQSGWNECASNHGCS 901
Qy 915 QLCIAIP-GGHRCCGASHYTLDPSSNCSPTTFLFSQKSAISRMIPDDQHSPLILPL 973
Db 902 HLCIAVFGVGCPCAHYSLNADNRCTASPTTFLFSQKSAINRMVWIDEQSPDILPI 961
Qy 974 HGLRNVKALDYDLDKFIYWDGGRON-IKRAKDGTQPP-VLTSLSQGNPDPRPHLSI 1031
Db 962 HSLRNVKALDYDLDKQLYWIDSRQNMIRKAQEDGSGQFTVVVSSVPSQNLIEIOPYLSI 1021
Qy 1032 DIYSRTLFTWCEATNTNVHRLSGEAMGVILRGDRDKPRAIVVNAERGILYFTNNQDRAA 1091
Db 1022 DIYSRYIYWTCEATNVINVTELDGRSVGLVKGQDRPRAIVVNPKEGYMYFTNLQERSP 1081
Qy 1092 KIERAALDGTBRVFTTGLIRPVALVVDNTLGLKLFWVDADLKRIESCDSLGSANRLTLED 1151
Db 1082 KIERAALDGTBRVLFPSGLSKPIALALDSRLGLFWADSLRLRIESSDLSGANRIVLED 1141
Qy 1152 ANIVQPLGLTLGKLYWIDRQOQMIERVEKTTGDKRTIOGRVAHLTGIIHAEVEVLEE 1211
Db 1142 SNILQPVGLTVFENWLYWIDRQOQMIKIDMTGREGRTKVQARQAQSLDIHAVKELNQE 1201
Qy 1212 FSAHPCARDNGCGSHICTAGDGTGPRSCPVHLVLLQNLITCGEPPTCSPOFACATGEI 1271
Db 1202 YRQHPCAQDNGCGSHICLVKGDGTGTRCSPHVLVLLQDELSCGEPPTCSPOQFTCTGEI 1261
Qy 1272 DCIFCAWRCDGFPBCDQSDDEEGCPVCSAAQPCARGQCVDLRLRCDEADCDQSDRSEAD 1331
Db 1262 DCIPVAMRCDFTECEDHSDDELNCPVCSQSFQPCASQCIDGALRCNGDANCQKSDEN 1321
Qy 1332 CDAICLPNQFCASQCQVLIKQOCDSPDCTDGSDELMECEITKPPSDSDSPAHSAGPVI 1391
Db 1322 CEVLCLIDQPCANGQCIGKHKCHDNVDCSDKDELDCYPTPEP---APQATNTVGSVI 1378
Qy 1392 GIILSLFVMGVYVFCORVVCORVAGANGPPFHEY-VSGTPHVPLNFTAPGCSQHPFTG 1450
Db 1379 GVIITVIFVSGTVYVFCORVVCORVAGANGPPFHEY-VSGTPHVPLNFTAPGCSQHPFTG 1438
Qy 1451 IACGKMSMSSVSLMGGRGVPLYDRNHVITGASSSSSSSTKATLYPPIINPPSPATDPSL 1510
Db 1439 MSRGKMSISSLINGGSSGPP-YDRAHVTGASSSSSSSTKGTYPPIINPPSPATDPSL 1497

Qy 1511 YNMDMFYSSNIPATAR--PYRPIIIRGMAPPPTPCSTDVCDSDYSASR-----WKASKYY 1563
Db 1498 YTFEFGYSSNSPSTHRSYRPSYRHFAPPTTFCSTDVCDSDYAPSRMTSVATAGYT 1557
Qy 1564 LDLSNDSDPVPPPTPHSQYLSAB---DSCPPSPATERSY-FHLPPPPSPCTDSS 1615
Db 1558 SDLNVDSEPVPPPTPRSQYLSABENYESCPPSPYTERSYSHHLYPPPPSPCTDSS 1613
RESULT 3
JE0273
low density lipoprotein receptor-related protein 6 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004
C:Accession: JE0273
R:Brown, S.D.; Twells, R.C.J.; Hey, P.J.; Cox, R.D.; Levy, E.R.; Soderman, A.R.; Metzker, Biochem. Biophys. Res. Commun. 248, 879-888, 1998
A:Title: Isolation and characterization of LRP6, a novel member of the low density lipoprotein receptor family
A:Reference number: JE0272; MUID:98369644; PMID:9704021
A:Accession: JE0273
A:Molecule type: mRNA
A:Residues: 1-1613 <BRO>
A:Cross-references: UNIPROT:O88572; GB:AF074265; NID:g3462528; PIDN:AAC33007.1; PID:g3462528
C:Genetics:
A:Gene: Lrp6
A:Map position: 6
C:Superfamily: low density lipoprotein receptor-related protein; EGF homology; LDL receptor
F:286-323/Domain: EGF homology <EGF1>
F:592-627/Domain: EGF homology <EGF2>
F:1207-1243/Domain: EGF homology <EGF3>
F:1249-1285/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:1288-1322/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:1326-1360/Domain: LDL receptor ligand-binding repeat homology <LDL3>
Query Match 69.8%; Score 6097.5; DB 2; Length 1613;
Best Local Similarity 69.3%; Pred. No. 0;
Matches 1120; Conservative 216; Mismatches 259; Indels 21; Gaps 12;
Qy 15 LLLLLLALCGCPAPAAASPLLFANRRDRLVADAGVKLESTIVVSGLEDAADVDFQFSK 74
Db 4 VLSRLTACSCFVLLRAA-PLLLYANRRDLRLVADATNGKENATIVVGGLEDAAADVFEHG 62
Qy 75 GAVVTVDSSEAIKQTYLNTGAOVNVI-SGLVSPDGLACDWGKKLYWTDSETNRLEV 134
Db 63 GLIYWSVSEAIKRTBFNKS-ESQVNVVSGLLSPDGLACDWGKKLYWTDSETNRLEV 121
Qy 135 ANLNGTSRKVLFWQDLQDQRAIALDPAHGYMYWTDWGETPRIERAGMDGSTRKIIVSDI 194
Db 122 SNLDGSLRKVLFWQDLQDQRAIALDPSGFGMYWTDWGEVPKIERAGMDGSRFVIINTEI 181
Qy 195 YWPNGLTIDLEBOQKLYWADAKLSFIHRANLDGSPRKVVEGSLTHPPALITLSDTLYWTD 254
Db 182 YWPNGLTIDYQERKLYWADAKLNIHKSNDGTNRQAVVKGSLPHFPFALTIFEDTLYWTD 241
Qy 255 WQTESIHACNKRCTGKKEILSALYSPMDIQVLSQERQPPFHTTRCEEDNGCASHLLCLSP 314
Db 242 WNTSHIACNKYTGEGUREIHSNIFSPMDIHAFQQQPNATNFCGIDNGGCSHLLCLMSP 301
Qy 315 SEPPYTCACPTGVOLQNGRTCKAGABEVLLAARLTDLRISLTDPTDFDITVLQVDDIRH 374
Db 302 VKPFYCACPTGVKLMENKTKDQATLLELLAARLTDLRISLTDPTDFDITVLQVDDIRH 361
Qy 375 AIAIDYDPEVGIYWTDDVEAIRRSFIDGSGSQFVVTQAIAHPDGIADVWVARNLYWTD 434
Db 362 AIAIDYDPEVGIYWTDDVEAIRRSFIDGSGSQFVVTQAIAHPDGIADVWVARNLYWTD 421
Qy 435 TGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVMGLMYWTDWGENPKICANLDGOER 494
Db 422 TGTDRIEVTRNGTWKILISEDLDEPRAIVLDPMVGMYWTDWGEIPKIERAALDGSR 481
Qy 495 RVLVNASLWPNGLALDQEGKLYWGDAKTDKIEVINVDGTRKRTLLEDKLPHPFGFTLL 554

Db 482 VVLVNTSLGWPNGLALDYDEGTIYWGDAKTDKIEVMNTDGTGRRVLVDEKIPHIFGFTLL 541
QY 555 GDFIYWTQWORSIERVHKVKASRDVVIDQLPDLMLGLKAVNVKVGTVNTPCADRNGGCSH 614
Db 542 GDFVYWTQWORSIERVHKRSAREVVIDQLPDLMLGLKATSVHVRIGSNPCAEADNGGCSH 601
QY 615 LCFEFTPHATRCGCPGLELLSMDKTCIPEAFVFTSRAAIIHRISETNNNDVAIPLTGV 674
Db 602 LCLYRPOGLRCACPIGFIIDMKTCIPEAFLLFSRRADIRISLETNNNNVAIPLTGV 661
QY 675 KEASALDFDVSNNHIYWTVDLSLKTISRPFMGSSVEHVFEGLDYPEGMVDMGKNLYW 734
Db 662 KEASALDFDVTNRIYWTVDLSLKTISRPFMGSSVEHVFEGLDYPEGMVDMGKNLYW 721
QY 735 ADTGNNRLEVARLDGQFVLYWRDLNDRSLALDPTKGYIYWTWGGKPRIVRAFMDGT 794
Db 722 ADTGNNRLEVSCLDQGHQVLYWKOLDSPRALDPAEGFMYWTWGGKPKIDRAAMDGS 781
QY 795 NCMTLVVDKVRANDLTIDYADQRLYWTDLDTNNIESSNNMLGQERVVIAADLPHPPGLTQY 854
Db 782 BRTTLVNVGRANGLTIDYAKERLYWTDLDTNLISSDMLGLNREVIADLPHPPGLTQY 841
QY 855 SYIYVTDNLHLSIERAKTSGRNTLIQGHLDYFVMDILVFHSSRODGLNCHMNNQCG 914
Db 842 QDIYVTDWRSRISIERANKTSQNRITIQGHLDYVMDILVFHSSRQAGWNECASSNGHCS 901
QY 915 QLCALIP-GCHRCGCASHYTLDPSSRNCSPPTTFLFSQKSAISRMIFDDQHSPLLIPL 973
Db 902 HCLAVPVGFGVCGCPAHYSLNADNRCTSAPEFLFLFSQKSAINRMVDEQOSPDIILPI 961
QY 974 HGLRNVKAIDYDPLDKFIYVNDGRQN-IKRAKMDGTQPF-VLTLSQCGNPDROPHDLSI 1031
Db 962 HSLRNVRAIDYDPLDKQLYWIDSRQNSIRKAHEDGGQGFNVVANSVANQNLIEQYDLSI 1021
QY 1032 DIYSRTLFWTCBATNTINVHRLSGSAMGVVLGRDRKPRAIVVNAERGLYFTNNQDRAA 1091
Db 1022 DIYSRIYWTCEATNVIDTRLDRSGVGLKGEQDRPRAIVVNEKGYMTFNLQERSP 1081
QY 1092 KIERAALDGTREVLFTGLIRPVALVVDNTLGLFWVDADLKRTESCDLSGANRLTLED 1151
Db 1082 KIERAALDGTREVLFTGLSKPIALALDLSKGLFWADSLRLRTESDLSGANRLVLED 1141
QY 1152 ANIVQPLGLITLGLKHLIYWDROQQMIERVEKTTGDKRTRIQGRVAHLTGIIHAEVLSLEE 1211
Db 1142 SNILQPVGLTFENWLYWIDROQQMIEXIDMTGREGRTKVQARIAQLSDIIHAEVLSLEE 1201
QY 1212 FSAHPCARDNGGCSHICTAKGDTGTRCSPVHLVLLQNLITCGEPPTCSPQFACATGEI 1271
Db 1202 YRQHPCAQDNGGCSHICLVKGDGTTTCSPVHLVLLQDELSCGEPPTCSPQFCTGTGDI 1261
QY 1272 DCIPGAWRCDFPECCDDGDEGCPVCSAAQPPCARGOCVDLRLRCDEADCDQSDRSDEAD 1331
Db 1262 DCIPVAWRCDFTECEDHSDDELNCVCSSESQFQASGQCIDGALKRCNGDANCQDKSDKN 1321
QY 1332 CDAICLPNQFCASQCQVLIKQCCDFFDCIDGSDLMCEITKPPSDPSPAHSSAIGBVI 1391
Db 1322 CEVLCLIDQFCANGCQVGHKKCDHSDVCDSDRDELDCYIEEP--APQATVTVGSVI 1378
QY 1392 GILLSLFVGGVYFVQVRVQVRVAGANGPPHPEYVSGTP-HVPLNFTIAPGSGHQPTG 1450
Db 1379 GVIVTIFVSGYIYFTICRMCLPRMKGDGETTNDYVHVSASVPLGYVPHPSLSGSLPG 1438
QY 1451 IACGKSMSSVSLMGGRGVPLYDRNHVTVGASSSSSTKATLYPPIINPPSPATDPSL 1510
Db 1439 MSRGKSMISSLSIMGGSGGP-YDRAHVTGASSSSSTKATLYPPIINPPSPATERSH 1497
QY 1511 YNMDVFGSNIPATAR--PYRPIYIRGMAPPTTCTDVCDDSDYSASR-----WKASKY 1563
Db 1498 YTMERFGYSSNPSHRSYXRPYSYRHFAPPTTCTDVCDDSDYAPSRMTSVATAGYT 1557
QY 1564 LDLNSDSDYPPTTPHQQYLSAE---DSCPPSPATERSY-FHLFPPPPSPCTDSS 1615
Db 1558 SDVNYDSEFVPPPTPRSQYLSAEENYSCPPSPYTERSYSHLYPPPPSPCTDSS 1613

RESULT 4
JB0373

Low density lipoprotein receptor related protein - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004

C:Accession: JB0373

R:Kim, D.; Inagaki, Y.; Suzuki, T.; Ioka, R.X.; Yoshioka, S.Z.; Magoori, K.; Kang, M.; Ch

J. Biochem. 124, 1072-1076, 1998

A:Title: A new low density lipoprotein receptor related protein, LRP5, is expressed in he

A:Reference number: JB0372; MUID:99054722; PMID:9832610

A:Accession: JB0373

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-527 <KIN>

A:Cross-references: UNIPROT:O77501; DBJ:AB017499

C:Superfamily: low density lipoprotein receptor-related protein; EGF homology; LDL recept

F:211-247/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F:250-284/Domain: LDL receptor ligand-binding repeat homology <LDL2>

F:288-322/Domain: LDL receptor ligand-binding repeat homology <LDL3>

Query Match 30.4%; Score 2654; DB 2; Length 527;

Best Local Similarity 91.6%; Pred. No. 1.1e-159;

Matches 482; Conservative 19; Mismatches 25; Indels 0; Gaps 0;

QY 1049 NVHRLSGSAMGVVLGRDRKPRAIVVNAERGLYFTNNQDRAAKIERAALDGTREVLFT 1108

Db 1 NVQRLSGEAVGVVLGRDRPRAIVVNAERGLYFTNNQDRAAKIERAALDGTREVLFT 60

QY 1109 TGLIRPVALVVDNTLGLFWVDADLKRTESCDLSGANRLTLEDANIYQPLGITLGLKHL 1168

Db 61 TGLIRPVALVVDNAGLKLFWVDADLKRTESCDLSGANRLTLEDASIVQPVGLVLRHLY 120

QY 1169 WIDRQQMIERVEKTTGDKRTRIQGRVAHLTGIIHAEVLSLEEFSAHPCARDNGGCSHIC 1228

Db 121 WADRQQMIERVEKATGDKRTVQGRVAHLTGIIHAEVLSLEEFSAHPCARDNGGCSHIC 180

QY 1229 IAKGDTGTRCSPVHLVLLQNLITCGEPPTCSPQFACATGEIDICIPGAWRCDFPECCDD 1288

Db 181 IAKGDTGTRCSPVHLVLLQNLITCAEPPTCSPQFACATGEIDICIPGAWRCDFPECCDD 240

QY 1289 QSDREGCPVCSAAQPPCARGOCVDLRLRCDEADCDQSDRSDEADCDALCLPNQFCASQC 1348

Db 241 QSDREGCPVCSAAQPPCARGOCVDLRLRCDEADCDHSDHSDADCDALCLPNQFCASQC 300

QY 1349 VLIQCCDFFDCIDGSDLMCEITKPPSDPSPAHSSAIGBVIIGILSLFVGGVYFVQ 1408

Db 301 LLIRQCCDFFDCVGDSDLMCEITRPFVDDIPAHSSAIGBVIIGILSLFVGGVYFVQ 360

QY 1409 RVVQVRVAGANGPPHPEYVSGTPHVPINFTIAPGSGHQPTGIACGKSMSSVSLMGGR 1468

Db 361 RVVQVRVAGANGPPHPEYVSGTPHVPINFTIAPGSGHQPTGIACGKSMSSVSLMGGR 420

QY 1469 GVPLYDRNHVTVGASSSSSTKATLYPPIINPPSPATDPSLYNMDVFGSNIPATAPY 1528

Db 421 AVPLIERNHVTGASSSSSTKASLYFQILNPPSPATDPSLYNMDVFGSNIPATAPY 480

QY 1529 RPYIIRGMAPPTTCTDVCDDSDYSASRWKASKYLLDLSDDSDYP 1574

Db 481 RSVYIRGMAPPTTCTDVCDDSDYSASRWKASKYLLDLSDDSDYP 526

RESULT 5
S02392

alpha-2-macroglobulin receptor precursor - human

N:Alternate names: CD91; LDL receptor-related protein 1; low density lipoprotein receptor

C:Species: Homo sapiens (man)

C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004

C:Accession: S02392; S30207; I37998; A39210; S42538

R:Herz, J.; Hamann, U.; Rognan, S.; Myklebost, O.; Gausepohl, H.; Stanley, K.K.

EMBO J. 7, 4119-4127, 1988

A:Title: Surface location and high affinity for calcium of a 500-kd liver membrane protei

A:Reference number: S02392; MUID:89210795; PMID:3266596
 A:Accession: S02392
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-4544 <HR>
 A:Cross-references: UNIPROT:Q07954; EMBL:X13916; NID:G34338; PIDN:CAA32112.1; PID:G34339
 R:Kristensen, T.
 submitted to the EMBL Data Library, October 1990
 A:Reference number: S30027
 A:Accession: S30027
 A:Molecule type: mRNA
 A:Residues: 3275-3864 <KRI>
 A:Cross-references: EMBL:X55077
 R:Herz, J.; Kowal, R.C.; Goldstein, J.L.; Brown, M.S.
 EMBO J. 9, 1769-1776, 1990
 A:Title: Proteolytic processing of the 600 kd low density lipoprotein receptor-related protein
 A:Reference number: S12538; MUID:90269210; PMID:2112085
 A:Contents: annotation; site of proteolytic cleavage
 R:Kutt, H.; Herz, J.; Stanley, K.K.
 Biochim. Biophys. Acta 1009, 229-236, 1989
 A:Title: Structure of the low-density lipoprotein receptor-related protein (LRP) promote
 A:Reference number: I37998; MUID:90089395; PMID:2597675
 A:Accession: I37998
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-11 <RES>
 A:Cross-references: EMBL:X15424; NID:G34408; PIDN:CAA33464.1; PID:G34409
 R:Strickland, D.K.; Ashcom, J.D.; Williams, S.; Burgess, W.H.; Migliorini, M.; Argraves, J. Biol. Chem. 265, 17401-17404, 1990
 A:Title: Sequence identity between the alpha2-macroglobulin receptor and low density lip
 A:Reference number: A39210; MUID:91009181; PMID:1698775
 A:Accession: A39210
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 150-166;234-338, 'X', 240-245, 'X', 247-252; 'G', 686-695;902-916;1096-1109; 'S', 17
 C:Genetics:
 A:Gene: GDB:LRP1; APR; LRP; A2MR
 A:Cross-references: GDB:119694; OMIM:107770
 A:Map position: 12q13.1-12q13.3
 C:Complex: The alpha-2-macroglobulin receptor complex consists of noncovalently-associated protein (see PIR:A39875).
 C:Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding
 C:Keywords: beta-hydroxyasparagine; beta-hydroxyaspartic acid; calcium binding; glycoprotein
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-3943/Product: alpha-2-macroglobulin receptor 515K chain #status predicted <515K>
 F:27-64/Domain: LDL receptor ligand-binding repeat homology <LDL1>
 F:72-108/Domain: LDL receptor ligand-binding repeat homology <LDL2>
 F:115-148/Domain: EGF homology <EG1>
 F:154-188/Domain: EGF homology <EG2>
 F:198-239/Domain: LDL receptor WYTD-containing repeat homology <YW01>
 F:240-281/Domain: LDL receptor WYTD-containing repeat homology <YW02>
 F:292-334/Domain: LDL receptor WYTD-containing repeat homology <YW03>
 F:335-378/Domain: LDL receptor WYTD-containing repeat homology <YW04>
 F:379-420/Domain: LDL receptor WYTD-containing repeat homology <YW05>
 F:421-468/Domain: LDL receptor WYTD-containing repeat homology <YW06>
 F:478-519/Domain: EGF homology <EG3>
 F:571-613/Domain: LDL receptor WYTD-containing repeat homology <YW07>
 F:614-659/Domain: LDL receptor WYTD-containing repeat homology <YW08>
 F:660-710/Domain: LDL receptor WYTD-containing repeat homology <YW09>
 F:711-752/Domain: LDL receptor WYTD-containing repeat homology <YW10>
 F:753-799/Domain: LDL receptor WYTD-containing repeat homology <YW11>
 F:807-842/Domain: EGF homology <EG4>
 F:854-890/Domain: LDL receptor ligand-binding repeat homology <LDL3>
 F:895-931/Domain: LDL receptor ligand-binding repeat homology <LDL4>
 F:936-971/Domain: LDL receptor ligand-binding repeat homology <LDL5>
 F:976-1011/Domain: LDL receptor ligand-binding repeat homology <LDL6>
 F:1015-1051/Domain: LDL receptor ligand-binding repeat homology <LDL7>
 F:1062-1097/Domain: LDL receptor ligand-binding repeat homology <LDL8>
 F:1104-1140/Domain: LDL receptor ligand-binding repeat homology <LDL9>
 F:1145-1182/Domain: LDL receptor ligand-binding repeat homology <LDL10>
 F:1185-1221/Domain: EGF homology <EG5>
 F:1227-1261/Domain: EGF homology <EG6>
 F:1269-1308/Domain: LDL receptor WYTD-containing repeat homology <YW12>

F:1309-1355/Domain: LDL receptor WYTD-containing repeat homology <YW13>
 F:1356-1398/Domain: LDL receptor WYTD-containing repeat homology <YW14>
 F:1399-1445/Domain: LDL receptor WYTD-containing repeat homology <YW15>
 F:1446-1488/Domain: LDL receptor WYTD-containing repeat homology <YW16>
 F:1489-1531/Domain: LDL receptor WYTD-containing repeat homology <YW17>
 F:1540-1578/Domain: EGF homology <EG7>
 F:1583-1626/Domain: LDL receptor WYTD-containing repeat homology <YW18>
 F:1627-1669/Domain: LDL receptor WYTD-containing repeat homology <YW19>
 F:1670-1713/Domain: LDL receptor WYTD-containing repeat homology <YW20>
 F:1714-1753/Domain: LDL receptor WYTD-containing repeat homology <YW21>
 F:1754-1796/Domain: LDL receptor WYTD-containing repeat homology <YW22>
 F:1797-1846/Domain: LDL receptor WYTD-containing repeat homology <YW23>
 F:1850-1886/Domain: EGF homology <EG8>
 F:1934-1976/Domain: LDL receptor WYTD-containing repeat homology <YW24>
 F:1977-2019/Domain: LDL receptor WYTD-containing repeat homology <YW25>
 F:2020-2063/Domain: LDL receptor WYTD-containing repeat homology <YW26>
 F:2064-2105/Domain: LDL receptor WYTD-containing repeat homology <YW27>
 F:2106-2151/Domain: LDL receptor WYTD-containing repeat homology <YW28>
 F:2159-2194/Domain: EGF homology <EG9>
 F:2199-2241/Domain: LDL receptor WYTD-containing repeat homology <YW29>
 F:2253-2294/Domain: LDL receptor WYTD-containing repeat homology <YW30>
 F:2344-2388/Domain: LDL receptor WYTD-containing repeat homology <YW31>
 F:2389-2429/Domain: LDL receptor WYTD-containing repeat homology <YW32>
 F:2430-2473/Domain: LDL receptor WYTD-containing repeat homology <YW33>
 F:2482-2517/Domain: EGF homology <EG10>
 F:2524-2561/Domain: LDL receptor ligand-binding repeat homology <LDL1>
 F:2566-2600/Domain: LDL receptor ligand-binding repeat homology <LDL2>
 F:2605-2639/Domain: LDL receptor ligand-binding repeat homology <LDL3>
 F:2652-2688/Domain: LDL receptor ligand-binding repeat homology <LDL4>
 F:2696-2730/Domain: LDL receptor ligand-binding repeat homology <LDL5>
 F:2734-2769/Domain: LDL receptor ligand-binding repeat homology <LDL6>
 F:2774-2812/Domain: LDL receptor ligand-binding repeat homology <LDL7>
 F:2818-2853/Domain: LDL receptor ligand-binding repeat homology <LDL8>
 F:2858-2897/Domain: LDL receptor ligand-binding repeat homology <LDL9>
 F:2904-2939/Domain: LDL receptor ligand-binding repeat homology <LDL10>
 F:2944-2980/Domain: EGF homology <EG11>
 F:2986-3021/Domain: EGF homology <EG12>
 F:3029-3068/Domain: LDL receptor WYTD-containing repeat homology <YW34>
 F:3069-3113/Domain: LDL receptor WYTD-containing repeat homology <YW35>
 F:3114-3156/Domain: LDL receptor WYTD-containing repeat homology <YW36>
 F:3157-3200/Domain: LDL receptor WYTD-containing repeat homology <YW37>
 F:3201-3241/Domain: LDL receptor WYTD-containing repeat homology <YW38>
 F:3242-3284/Domain: LDL receptor WYTD-containing repeat homology <YW39>
 F:3294-3330/Domain: EGF homology <EG13>
 F:3334-3369/Domain: LDL receptor ligand-binding repeat homology <LDL11>
 F:3374-3408/Domain: LDL receptor ligand-binding repeat homology <LDL12>
 F:3413-3448/Domain: LDL receptor ligand-binding repeat homology <LDL13>
 F:3453-3489/Domain: LDL receptor ligand-binding repeat homology <LDL14>
 F:3494-3531/Domain: LDL receptor ligand-binding repeat homology <LDL15>
 F:3536-3570/Domain: LDL receptor ligand-binding repeat homology <LDL16>
 F:3575-3609/Domain: LDL receptor ligand-binding repeat homology <LDL17>
 F:3613-3647/Domain: LDL receptor ligand-binding repeat homology <LDL18>
 F:3654-3690/Domain: LDL receptor ligand-binding repeat homology <LDL19>
 F:3695-3731/Domain: LDL receptor ligand-binding repeat homology <LDL20>
 F:3741-3776/Domain: LDL receptor ligand-binding repeat homology <LDL21>
 F:3785-3822/Domain: EGF homology <EG14>
 F:3828-3860/Domain: EGF homology <EG15>
 F:3868-3911/Domain: LDL receptor WYTD-containing repeat homology <YW40>
 F:3912-3959/Domain: LDL receptor WYTD-containing repeat homology <YW41>
 F:3944-4420/Product: alpha-2-macroglobulin receptor 85K chain #status predicted <85K>
 F:3944-4420/Domain: 85K chain extracellular #status predicted <EXT>
 F:3970-4012/Domain: LDL receptor WYTD-containing repeat homology <YW42>
 F:4013-4056/Domain: LDL receptor WYTD-containing repeat homology <YW43>
 F:4057-4099/Domain: LDL receptor WYTD-containing repeat homology <YW44>
 F:4100-4142/Domain: LDL receptor WYTD-containing repeat homology <YW45>
 F:4151-4182/Domain: EGF homology <EG16>
 F:4200-4231/Domain: EGF homology <EG17>
 F:4236-4267/Domain: EGF homology <EG18>
 F:4272-4303/Domain: EGF homology <EG19>
 F:4308-4339/Domain: EGF homology <EG20>
 F:4344-4374/Domain: EGF homology <EG21>
 F:4377-4408/Domain: EGF homology <EG22>
 F:4421-4444/Domain: transmembrane #status predicted <TM>

F:4445-4544/Domain: intracellular #status predicted <INT>
 F:166,2998/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
 F:2958/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
 F:4075,4125,4278/Binding site: carbonylate (Asn) (covalent) #status predicted

Query Match: 25.1%; Score 2193; DB 1; Length 4544;
 Best Local Similarity 35.1%; Pred. No. 3,1e-129;
 Matches 521; Conservative 266; Mismatches 575; Indels 124; Gaps 54;

QY 33 PLLLFANRDRVLVDAGGVKLESTTVVSGLEDAANVDFQFSGAVYWTVDVSEEAIKQYVL 92
 DB 1269 PFIIFSNHEIRRIDLH--KGDSYLVPLGLRNTIALDFHLSQSALYWTVDVVEDKIYRGKL 1326

QY 93 NQTGA--AVONVVISGLVSPDGLACDWGKKLYWTDSETRNLEVANLNGTSRKLFWODL 150
 DB 1327 LONGALTSEFVVIQGLATPEGLADVNIAGNIYVESNLDQIEVAKLDTGRTTLTLLAGDI 1386

QY 151 DQPRALDPAGHYMYTDM-GETPRIERAGMDGSTRKII--VDSDIYWPNGLTITDLREQ 207
 DB 1387 EHPRAIALDPDGLFTWDASLPLRIEASMSGAGRTVHRETSGCGWPNGLTVDYLEK 1446

QY 208 KLYWADAKLSFIHRANLDGSRQKVVES--LTHPFALTSGDTLYWTDWQTRSIHANCK 265
 DB 1447 RILWTDARSDAIYSARYDGSCHMEVLRGHEFLSHFPFVLYXGEVYWTDRNTTLAKANK 1506

QY 266 RTGGRKEILSALYSPMDIQVLQSERQPFPHTRCEDNGG--CSHLCLLSPSEPFYTC 322
 DB 1507 WGHNVTVQRTNTQPFQVTHPSRQPMANPCEA--NGGQGPCSHLCLINRT-VSCA 1564

QY 323 CPTGVQLDNGRTKAGAEVLLARRTDLRRISLDTDPDFTDIV--LQVDDIRHAIADYD 381
 DB 1565 CPHLMKLHKDNTTCVE--FKKFLYARQMEIRGVLDAPYVNYIISFTVPDIDNVTLVDYD 1623

QY 382 PLEGVYWTDEVARIRAYLDGSAQTLVNTENPDGIVADVARNLYWT--DTGTDR 439
 DB 1624 AREQRYVMSDVRTQAKRAFTNGTGVETVSVSADLPNAHGLAVDWVSRNLFWTSTYDINKKQ 1683

QY 440 LEVTRLNGTSKILVSEDLDEPRALHPVLMGLMYTWDGNGENPKIECANLQSGERRVLVN 499
 DB 1684 INVARKDSFKNVY--QGLEQPHGLVHPHRLGKLYWD--GN--ISMAMDGSNRTLFS 1739

QY 500 ASLGNPNGLALDQEGKLYWGDAKTDKIEVINVDGTRKRTL--LEDKLPHFIFGFTLLGDF 557
 DB 1740 GQKG--PVGLAIDFPESKLYWISSGNHTINRCNLDSGLEVIDAMRSQLGKATALAIMGDK 1798

QY 558 IYWTWQRRSIRERHKVKAARDVIIDQLPDL--MGLKAVNVAKV---GTNFCARDNGCS 613
 DB 1799 LWNADQVSEKMGTCADGSGSVVLRSNTTILVHMVKVYDESQILDHKGNTPCSVNNGDCS 1858

QY 614 HLCFFTPHATR--CGCPIGLELLSDMKTCTVPEAFVFTSRAAIHRIISLETNN--NDVAIPL 671
 DB 1859 QLCLPTSETTSQCTAGYSLSRSGQACGVSFLYSVHGGIRGIPDPNDKSDALVPV 1918

QY 672 TGVKEASALDFVNNHIIYWTVDLSLTKTISRPFMGSSVEHVVEFGLDYPEGMAVDWGMKN 731
 DB 1919 SGTSLAVGIDFHAENDTIYVDMGLSTISRAKRDQTWREDDVVTNGIGRVEGIADVIAIGN 1978

QY 732 LYWADTGNIRLEVARLDQFQVLRWRLDNPSRLADPTKYGIYWTWGWGKPIRVAFPM 791
 DB 1979 IYWTDDQGDVIEVARLNGSFYVWISQGLDPRAITVHPPEKGYLFTWEGQYPIERSRL 2038

QY 792 DGTNMTLVD--KVGRLNDLTDYADORLYWTDLTNMTESNN--LQGBR--VVIADDLPH 848
 DB 2039 DGTERRVLNVYSIWPNGISVDYQDKLYWCARDTKLERIDLETGENREVVLSNNMDM 2098

QY 849 FGLTQYSYIYWTWNLH--SIERADKTSGRNRTLIQCHLDP--VMDILVFHSSRQDGLN 904
 DB 2099 FSVSVFEDFIYWD--RTHANGSIKSGKDNATDSVPLRTGIGVQLKDIKVENRDRQKGTN 2157

QY 905 DCMENHNGCGCLAIPOGHR--CCASHYTLDPSSRNCSPPTTFLLSQKSAISRM--IPD 962
 DB 2158 VCAVANGGCGQLYRGQRQACACA--HGMLEADGASCREYAGYLLYSERTILKSIHLSD 2216

QY 963 DOHSPDLILPLHG---LNNKVAIDVD-----PLDKFIYVDGR-ONIKRAKDDGTOPF 1011
 DB 2217 ERNLNAPVQPPDEDEHMKVNIALAFDYAGTSPGTPNRIFFSDIHFGNIQIINDGSRRI 2276

QY 1012 VLTSLSQSQNDPDRPHLSDID--IYSR---TLFWTCEATNTINVHRLSGEAMG----- 1059
 DB 2277 TIV-----ENVG-----SVEGLAYHGRWDLYWTSYTTSTIIRHTVDTQTPGAFERTV 2325

QY 1060 VLRGDRDKPRAIVVNAERGVLTYTNMQDRAAKIERAALDGTREVLFTTGLIRVALVV 1119
 DB 2326 ITMSGD-DHPRAFVLDBCCQNLMTWNNEQHPSIMRAALSGANVLTLEKDIRTPENGLAI 2384

QY 1120 DNTLGKLFVWDADLKRIESCDSLGSANRLTLEDANIVOPGLTILGKHLTYWIDRQOQMIER 1179
 DB 2385 DHRAEKLFSRATLDKERCYDGHSHRVILKSPVHPFGLAVYGEHIFWTDWVRVQR 2444

QY 1180 VEKTTGKRTIRIQGRVAHL--TGIHAV--EEVSELEFSAHPCARDNGGCSHICIAKGDGTPR 1237
 DB 2445 ANKHVGSNMKLRLVDIPOQPMGIIAVANDTNSCELS--PCRINNGCQDCLLTTHQGHVN 2502

QY 1238 CSCPHVLVLLQNLITC-GEPTTC--SPQOPACATGEIDCIPGAWRCDFGPECDQSDSEB-- 1293
 DB 2503 CSCRGRI--LQDDLTCRAVNSSCRAQDEFECANGE--CINFSLTCDGVPHCKSKSDEKPS 2559

QY 1294 --GCPVCSAAQFPARGQCVDLRLRCDEADCDRDEADCD-AICLPNQCASGQCVL 1350
 DB 2560 YCNSRCKTKTRQCSNGRCVSNMLWCNGADDCGDSDEIPCNKTACGVGEFRCDGTGIG 2619

QY 1351 IKQCDSPDCIDGSDLMCEITKPPSDSDSPAHSAGIPVIGIILSLFPMVGVPVFCQV 1410
 DB 2620 NSSRCNQFVDCEDASDEMNSATDCSS-----YFRLGVKGVLPQCERT 2663

QY 1411 -----VCORYAGAN--GPFPHEY-VSGT--PHVPLNFIA-PGG 1442
 DB 2664 SLYAPSWVCD---GANDCGDYSDERDCPGVKRPRCLNYPACPSG 2706

RESULT 6
 A53102
 alpha-2-macroglobulin receptor precursor - chicken
 N:Alternate names: CP91; LDL receptor-related protein 1; low density lipoprotein receptor
 C:Species: Gallus gallus (chicken)
 C:Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 09-Jul-2004
 C:Accession: A53102
 R:Nimf, J.; Stifani, S.; Bilous, P.T.; Schneider, W.J.
 J. Biol. Chem. 269, 212-219, 1994
 A:Title: The somatic cell-specific low density lipoprotein receptor-related protein of the chicken
 A:Reference number: A53102; MUID:94103212; PMID:7506255
 A:Accession: A53102
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-4543 <NIM>
 A:Cross-references: UNIPROT:P98157; GB:X74904; NID:9438006; PIDN:CA52870.1; PID:9438007
 C:Complex: The alpha-2-macroglobulin receptor complex consists of noncovalently-associated protein.
 C:Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding
 C:Keywords: beta-hydroxyasparagine; beta-hydroxyaspartic acid; calcium binding; glycoprotein
 F:i-17/Domain: signal sequence #status predicted <SIG>
 F:18-3942/Domain: alpha-2-macroglobulin receptor 515K chain #status predicted <515K>
 F:18-3942,3943-4543/Product: alpha-2-macroglobulin receptor #status predicted <MAT>
 F:29-66/Domain: LDL receptor ligand-binding repeat homology <LDL1>
 F:74-110/Domain: LDL receptor ligand-binding repeat homology <LDL2>
 F:117-150/Domain: EGF homology <EG1>
 F:156-190/Domain: EGF homology <EG2>
 F:200-241/Domain: LDL receptor YWTD-containing repeat homology <YW01>
 F:242-283/Domain: LDL receptor YWTD-containing repeat homology <YW02>
 F:294-336/Domain: LDL receptor YWTD-containing repeat homology <YW03>
 F:337-380/Domain: LDL receptor YWTD-containing repeat homology <YW04>
 F:381-422/Domain: LDL receptor YWTD-containing repeat homology <YW05>
 F:423-470/Domain: LDL receptor YWTD-containing repeat homology <YW06>
 F:480-521/Domain: EGF homology <EG3>
 F:573-615/Domain: LDL receptor YWTD-containing repeat homology <YW07>
 F:616-661/Domain: LDL receptor YWTD-containing repeat homology <YW08>

F:662-712/Domain: LDL receptor YWTD-containing repeat homology <YW09>
F:713-754/Domain: LDL receptor YWTD-containing repeat homology <YW10>
F:755-797/Domain: LDL receptor YWTD-containing repeat homology <YW11>
F:805-840/Domain: EGF homology <EG4>
F:852-888/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:893-929/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:934-969/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F:974-1009/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:1013-1049/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F:1060-1095/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F:1102-1138/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:1143-1180/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:1183-1219/Domain: EGF homology <EG5>
F:1225-1259/Domain: EGF homology <EG6>
F:1267-1306/Domain: LDL receptor YWTD-containing repeat homology <YW12>
F:1307-1353/Domain: LDL receptor YWTD-containing repeat homology <YW13>
F:1354-1396/Domain: LDL receptor YWTD-containing repeat homology <YW14>
F:1397-1443/Domain: LDL receptor YWTD-containing repeat homology <YW15>
F:1444-1486/Domain: LDL receptor YWTD-containing repeat homology <YW16>
F:1487-1529/Domain: LDL receptor YWTD-containing repeat homology <YW17>
F:1538-1576/Domain: EGF homology <EG7>
F:1581-1624/Domain: LDL receptor YWTD-containing repeat homology <YW18>
F:1625-1667/Domain: LDL receptor YWTD-containing repeat homology <YW19>
F:1668-1711/Domain: LDL receptor YWTD-containing repeat homology <YW20>
F:1712-1751/Domain: LDL receptor YWTD-containing repeat homology <YW21>
F:1752-1794/Domain: LDL receptor YWTD-containing repeat homology <YW22>
F:1795-1842/Domain: LDL receptor YWTD-containing repeat homology <YW23>
F:1846-1882/Domain: EGF homology <EG8>
F:1930-1972/Domain: LDL receptor YWTD-containing repeat homology <YW24>
F:1973-2015/Domain: LDL receptor YWTD-containing repeat homology <YW25>
F:2016-2059/Domain: LDL receptor YWTD-containing repeat homology <YW26>
F:2060-2101/Domain: LDL receptor YWTD-containing repeat homology <YW27>
F:2102-2147/Domain: LDL receptor YWTD-containing repeat homology <YW28>
F:2155-2190/Domain: EGF homology <EG9>
F:2195-2237/Domain: LDL receptor YWTD-containing repeat homology <YW29>
F:2247-2288/Domain: LDL receptor YWTD-containing repeat homology <YW30>
F:2338-2382/Domain: LDL receptor YWTD-containing repeat homology <YW31>
F:2383-2423/Domain: LDL receptor YWTD-containing repeat homology <YW32>
F:2424-2467/Domain: LDL receptor YWTD-containing repeat homology <YW33>
F:2476-2511/Domain: EGF homology <EG10>
F:2518-2555/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:2560-2594/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:2599-2633/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:2646-2682/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:2690-2724/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:2732-2767/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:2772-2810/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:2816-2851/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:2856-2895/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:2902-2936/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:2941-2977/Domain: EGF homology <EG11>
F:2983-3018/Domain: EGF homology <EG12>
F:3026-3065/Domain: LDL receptor YWTD-containing repeat homology <YW34>
F:3066-3110/Domain: LDL receptor YWTD-containing repeat homology <YW35>
F:3111-3153/Domain: LDL receptor YWTD-containing repeat homology <YW36>
F:3154-3197/Domain: LDL receptor YWTD-containing repeat homology <YW37>
F:3198-3238/Domain: LDL receptor YWTD-containing repeat homology <YW38>
F:3239-3281/Domain: LDL receptor YWTD-containing repeat homology <YW39>
F:3291-3327/Domain: EGF homology <EG13>
F:3331-3366/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:3371-3405/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:3410-3445/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:3450-3486/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:3491-3528/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:3533-3567/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:3572-3606/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:3610-3644/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:3651-3687/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:3692-3728/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:3738-3774/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:3783-3820/Domain: EGF homology <EG14>
F:3826-3858/Domain: EGF homology <EG15>
F:3866-3909/Domain: LDL receptor YWTD-containing repeat homology <YW40>

F:3910-3968/Domain: LDL receptor YWTD-containing repeat homology <YW41>
F:3943-4543/Domain: alpha-2-macroglobulin receptor 85K chain #status predicted <85K>
F:3943-4420/Domain: 85K chain extracellular #status predicted <EXT>
F:3969-4011/Domain: LDL receptor YWTD-containing repeat homology <YW42>
F:4012-4055/Domain: LDL receptor YWTD-containing repeat homology <YW43>
F:4056-4098/Domain: LDL receptor YWTD-containing repeat homology <YW44>
F:4099-4141/Domain: LDL receptor YWTD-containing repeat homology <YW45>
F:4150-4181/Domain: EGF homology <EG16>
F:4199-4230/Domain: EGF homology <EG17>
F:4235-4266/Domain: EGF homology <EG18>
F:4271-4302/Domain: EGF homology <EG19>
F:4307-4338/Domain: EGF homology <EG20>
F:4343-4373/Domain: EGF homology <EG21>
F:4376-4408/Domain: EGF homology <EG22>
F:4421-4443/Domain: transmembrane #status predicted <TM>
F:4444-4543/Domain: intracellular #status predicted <INT>
F:116,138,387,276,359,448,731,926,1048,1152,1153,1193,1216,1305,1509,1556,1573,1614,1643,
3485,3659,3786,3837,3952,4074,4124,4178,4278/Binding site: carboxylate (Asn) (covalent)
F:168,2995/Modified site: erythro-beta-hydroxyaspartate (Asn) #status predicted
F:2955/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted

Query Match 24.9%; Score 2180; DB 1; Length 4543;

Best Local Similarity 34.0%; Pred. No. 2.1e-128; Indels 130; Gaps 49;
Matches 506; Conservative 270; Mismatches 581;

QY 33 PLLLFANRRDVRVLVADAGGVKLESTIVVSGLEDAADVDFQFSKGVAVYTDVSEAIKQTYL 92
DB 1267 PPIFSNHEIRIDLH--RGDYSVLVPLGENTLALDHLNQSSLYWTDVVDKIYRGKL 1324
QY 93 NQTGA--AVQNVITSLVSPDGLACDWVKKLYWTDSETRIEVANLNGTSRKLFWQDL 150
DB 1325 LENGATSFEEVQVIGLATPEGLAVDIAGNIYVWESNLDQIEVAKLDGTWRTTLAGDI 1384
QY 151 DQPAIALDPAGHYMYWTDW--GETPRIERAGMDGSTRKI--VDSDIYWPNGLTLDLREQ 207
DB 1385 EHPAIALDPRIYGLFTWTDWASLPRIEASMSGAGRTHHKTGSGGWPNGLTVVDYLEK 1444
QY 208 KLYWADAKLSFIHRANLDGSPQKQVWEGS--LTHPPFALTLSGDTLYWTDWOTRSIHACNK 265
DB 1445 RILWIDARSDAIYSLVDGTGHIIEVLRGHEVLSHPFVTLVGGVEYTDWRTNLAKANK 1504
QY 266 RTGGRKRKELISALYSPMDIQVLSQERQPFHTRCEENGG---CSHLCLLSPSEPFYCA 322
DB 1505 WTGHNVTVQRTNTQPPDLQVYHPSRQPLADNPCEA--NGKGKPCSHLCLINYNRTL--SCA 1562
QY 323 CPTGVQIQDNGRTCKAGAEVILLARTRDLRISLDTPDFDIV--LQVDDIRHAIADYD 381
DB 1563 CPHLMKLDKONTTCYE--FKFELLYARQMEIRGVDDINPYNIIISFTVPDIDNVTVVDYD 1621
QY 382 PLEGVYVWTDDEVARIRRAYLDGSGAOTLVNTEINDPGIADVDWVARNLYWT--DTGTR 439
DB 1622 AVEQRIYWSVTRTQIKRAFINGTVEIVSADLPNAHGLSVDWVSRNLFWTSYDTNKKQ 1681
QY 440 IEVTRLNGTSRKILVSEDLDEPRALALHPVNGMLYWTDWGNPKIECANLQDQERRVLVN 499
DB 1682 INVARLDG--SPKNAVIQGLDKPHCLVHPLHGLKLYWTD--GDN--ISVANMDGSRNLTLLFT 1737
QY 500 ASLQWPNGLALDQEGKLYWGDAXTDKIEVINVDGTRKRTL--LEDKLPHPFGFTLLGDF 557
DB 1738 NQRG--PVGLAIDYPEKLYWISSGNGTINRCLNLDGSDLEVTVAVKQSLSKATALAIMGDK 1796
QY 558 IYWTDMQRRSIRERVHVKASRDVILDQLPDLMLGLKAVNVAKV--VGTNPACDRNGCCSHL 615
DB 1797 LWNADQASERMGTCKKDGTEVTVLRNSTTLVLMKVVDESIQAGNSPCSVNNGDCQQL 1856
QY 616 CFFTPHATR--CCPGIGLELLSDMKTCIVPEAPLVFTSRAAHRISLETNN--NDVAIPITG 673
DB 1857 CLPTSETSRSCWCTAGYSLKSGQCEGVSFLLYSVHEGIRGIPLDPNKSDALVPVSG 1916
QY 674 VKESALDPVSNHIIYWTDVSLKTIISRAFWNGSVSEHVEFGLDYPSGMADVWNGKLY 733
DB 1917 TSLAVGIDFHAENDIYWDMGLSTISRKRDQTWREDVWTVNGIRGVEGIADVWVAGNIY 1976

```

QY 734 WADTCNRIEVARLDGQROVLVWRDLNPRSLALDPTKGYLYWTWGGKPRIVRAFDWG 793
Db 1977 WTDQGFVIEVARLNGSRVYVVISQGLDKPRAITVHPKGYLFWTEWQYPRIESRLDG 2036
QY 794 TNCMTLVD-KVGRANDLTADYDQRLYTDLDLNNIESNM-LQGER-VVIADLLPHFPFG 850
Db 2037 TERWLVNVSISWPNGISVYDYGKLYWCDAKTIERIDLETGENREVLSSDNWDFPS 2096
QY 851 LTQYSDIYIYTDWNLH----SIRADKTSGRNRTLQGHLDLF-VMDILYFHSRQGLNDC 906
Db 2097 VSVFEDIYIYMSD-RTHANGSIKRGSKDNATESVSLRTGIGVOLKDIKVENRARKGTNVV 2155
QY 907 MHNNGCGQLCIAIPGHR-CGCASHYTLDPSSRNCSPPTPLFSSQSAISRMIPDDQH 965
Db 2156 AQNNGCGQLCFRGGGRRTACA-HGLSESDGVSCRDYDGYLLYSERTILKSHLSDEN 2214
QY 966 SPDLILPL-----HGLRNKAIDYDPLDKFYWVDRGQ-----NFKRAXDDGT- 1008
Db 2215 --NLNAPIKPPEDAHEMKNVIALAFD---YRYGTGKSNRIFYSDIHFGNIQQINDDGTG 2268
QY 1009 QPVLTSLSQGNQPRQPHDLSID--IYSR---TLFWTCEATNTINVRHLSGEAMG---- 1059
Db 2269 RRTIVENVG-----SVEGLAYHRGWDTLYWTSYTTSTTTRHTVQSRILGAFER 2316
QY 1060 ---VVLGRDRKPRAI VVNAERGYLYFTNMODRAAKIERAALDGTREVLFTTGLIRPVA 1116
Db 2317 ETVITMSGD-DHPRAFVLDDECONLFWTNWNEQHPISINRATLSGANVLIIIDQIRTPNG 2375
QY 1117 LVVDNTLKLFWVDADLKRIESCDLSGANRLTLEDANIVQPLGLTILGKHLIYWDROOQM 1176
Db 2376 LAIDHRAEKIVFSDATLKDIERCEYDGGSHRVLKSEPHVPFGLAVGYDIYFTWDWVRA 2435
QY 1177 IERVEKTTGDKRTRIQGRVAHL-TGIHAVEVSVLEEPSAHCARDNGGCSHICTAKDGT 1235
Db 2436 VORANKYYGTDMKLRIDIPQOPMGI IIAVAN-DTDSCLSLSPCRVNVNGCQDLCLLTPKXG 2494
QY 1236 PRCSFVHLVLLQNLITCCEPPTCS-PQOFACATGEIDICPGAWRCDFPECDQSDDEE- 1293
Db 2495 VNCSCRGRLVQEDFTCKALNSTCNVHDEFEGNG--DCIDFSRTCDGVVHCKKSDSKQ 2552
QY 1294 ---GCPVCSAAQFFCARGCQVDRLRLRCDEAGDCQDRSDEADC-ALICLPNFCASGQCV 1349
Db 2553 SYCSRSKCKGFLHCHMGRCVASRFWCNGVDVDCGNSDEVPCNKTSCAATEFRCDRGSCI 2612
QY 1350 LIKQCDSPFCIDGSDLMCEITKPPSPDDSPAHSAITGPVIGILSLIFVWGGVYFVQCR 1409
Db 2613 GNSRSCNFIDCEDASDEMNCTATDCSS-----YFKLGVKGTTFQKCEH 2656
QY 1410 V-----VQORVAGAN--GPFPEHY---VSGTPHVPINFA-PGG 1442
Db 2657 TSLCVAPSWVCD---GANDCGDYSERNCPGGRKPKCANYFACPSG 2700

RESULT 7
alpha-2-macroglobulin receptor precursor - mouse
N:Alternate names: CD91; LDL receptor-related protein 1; low density lipoprotein receptor
C:Species: Mus musculus (house mouse)
C>Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 09-Jul-2004
C:Accession: S25111; S32554
R:van Leuven, F.
submitted to the EMBL Data Library, July 1992
A:Reference number: S25111
A:Accession: S25111
A:Molecule type: mRNA
A:Residues: 1-4545 <VANI>
A:Cross-references: UNIPROT:Q61291; EMBL:X67469; NID:g49941; PIDN:CAA47817.1; PID:g49942
R:van Leuven, F.; Stas, L.; Raymakers, L.; Overbergh, L.; de Strooper, B.; Hilliker, C.;
Biochim. Biophys. Acta 1173, 71-74, 1993
A:Title: Molecular cloning and sequencing of the murine alpha-2-macroglobulin receptor c
A:Reference number: S32554; MUID:93250049; PMID:8485155
A:Accession: S32554
A>Status: nucleic acid sequence not shown

```

```

A:Molecule type: mRNA
A:Residues: 1-28;4416-4453 <VANI>
A:Cross-references: EMBL:X67469
C:Complex: The alpha-2-macroglobulin receptor complex consists of noncovalently-associated protein (see PIR:J02811).
C:Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding
C:Keywords: beta-hydroxyasparagine; beta-hydroxyaspartic acid; calcium binding; glycoprotein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-3944,3945-4545/Product: alpha-2-macroglobulin receptor #status predicted <MAT>
F:20-3944/Domain: alpha-2-macroglobulin receptor 515K chain #status predicted <515K>
F:28-65/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:73-109/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:116-149/Domain: EGF homology <EG1>
F:155-189/Domain: EGF homology <EG2>
F:199-240/Domain: LDL receptor WYTD-containing repeat homology <YW01>
F:241-282/Domain: LDL receptor WYTD-containing repeat homology <YW02>
F:293-335/Domain: LDL receptor WYTD-containing repeat homology <YW03>
F:336-379/Domain: LDL receptor WYTD-containing repeat homology <YW04>
F:380-421/Domain: LDL receptor WYTD-containing repeat homology <YW05>
F:422-469/Domain: LDL receptor WYTD-containing repeat homology <YW06>
F:479-520/Domain: EGF homology <EG3>
F:572-614/Domain: LDL receptor WYTD-containing repeat homology <YW07>
F:615-660/Domain: LDL receptor WYTD-containing repeat homology <YW08>
F:661-711/Domain: LDL receptor WYTD-containing repeat homology <YW09>
F:712-753/Domain: LDL receptor WYTD-containing repeat homology <YW10>
F:754-800/Domain: LDL receptor WYTD-containing repeat homology <YW11>
F:808-843/Domain: EGF homology <EG4>
F:855-891/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:896-932/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:937-972/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F:977-1012/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:1016-1052/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F:1063-1098/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F:1105-1141/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:1146-1183/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F:1186-1222/Domain: EGF homology <EG5>
F:1228-1262/Domain: EGF homology <EG6>
F:1270-1309/Domain: LDL receptor WYTD-containing repeat homology <YW12>
F:1310-1356/Domain: LDL receptor WYTD-containing repeat homology <YW13>
F:1357-1399/Domain: LDL receptor WYTD-containing repeat homology <YW14>
F:1400-1446/Domain: LDL receptor WYTD-containing repeat homology <YW15>
F:1447-1489/Domain: LDL receptor WYTD-containing repeat homology <YW16>
F:1490-1532/Domain: LDL receptor WYTD-containing repeat homology <YW17>
F:1541-1579/Domain: EGF homology <EG7>
F:1584-1627/Domain: LDL receptor WYTD-containing repeat homology <YW18>
F:1628-1670/Domain: LDL receptor WYTD-containing repeat homology <YW19>
F:1671-1714/Domain: LDL receptor WYTD-containing repeat homology <YW20>
F:1715-1754/Domain: LDL receptor WYTD-containing repeat homology <YW21>
F:1755-1797/Domain: LDL receptor WYTD-containing repeat homology <YW22>
F:1798-1847/Domain: LDL receptor WYTD-containing repeat homology <YW23>
F:1851-1887/Domain: EGF homology <EG8>
F:1935-1977/Domain: LDL receptor WYTD-containing repeat homology <YW24>
F:1978-2020/Domain: LDL receptor WYTD-containing repeat homology <YW25>
F:2021-2064/Domain: LDL receptor WYTD-containing repeat homology <YW26>
F:2065-2106/Domain: LDL receptor WYTD-containing repeat homology <YW27>
F:2107-2152/Domain: LDL receptor WYTD-containing repeat homology <YW28>
F:2160-2195/Domain: EGF homology <EG9>
F:2200-2242/Domain: LDL receptor WYTD-containing repeat homology <YW29>
F:2254-2295/Domain: LDL receptor WYTD-containing repeat homology <YW30>
F:2345-2389/Domain: LDL receptor WYTD-containing repeat homology <YW31>
F:2390-2430/Domain: LDL receptor WYTD-containing repeat homology <YW32>
F:2431-2474/Domain: LDL receptor WYTD-containing repeat homology <YW33>
F:2483-2518/Domain: EGF homology <EG10>
F:2525-2562/Domain: LDL receptor ligand-binding repeat homology <LDL11>
F:2567-2601/Domain: LDL receptor ligand-binding repeat homology <LDL12>
F:2606-2640/Domain: LDL receptor ligand-binding repeat homology <LDL13>
F:2653-2689/Domain: LDL receptor ligand-binding repeat homology <LDL14>
F:2697-2731/Domain: LDL receptor ligand-binding repeat homology <LDL15>
F:2735-2770/Domain: LDL receptor ligand-binding repeat homology <LDL16>
F:2775-2813/Domain: LDL receptor ligand-binding repeat homology <LDL17>
F:2819-2854/Domain: LDL receptor ligand-binding repeat homology <LDL18>
F:2859-2898/Domain: LDL receptor ligand-binding repeat homology <LDL19>
F:2905-2940/Domain: LDL receptor ligand-binding repeat homology <LDL20>

```

F:2945-2981/Domain: EGF homology <EG11>
F:2987-3022/Domain: EGF homology <EG12>
F:3030-3069/Domain: LDL receptor YWTD-containing repeat homology <YW34>
F:3070-3114/Domain: LDL receptor YWTD-containing repeat homology <YW35>
F:3115-3157/Domain: LDL receptor YWTD-containing repeat homology <YW36>
F:3158-3201/Domain: LDL receptor YWTD-containing repeat homology <YW37>
F:3202-3242/Domain: LDL receptor YWTD-containing repeat homology <YW38>
F:3243-3285/Domain: LDL receptor YWTD-containing repeat homology <YW39>
F:3295-3331/Domain: EGF homology <EG13>
F:3335-3370/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:3375-3409/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:3414-3449/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:3454-3490/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:3495-3532/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F:3537-3571/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:3576-3610/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F:3614-3648/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F:3655-3691/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:3696-3732/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F:3742-3777/Domain: LDL receptor ligand-binding repeat homology <LDL11>
F:3786-3823/Domain: EGF homology <EG14>
F:3829-3861/Domain: EGF homology <EG15>
F:3869-3912/Domain: LDL receptor YWTD-containing repeat homology <YW40>
F:3913-3970/Domain: LDL receptor YWTD-containing repeat homology <YW41>
F:3945-4545/Domain: alpha-2-macroglobulin receptor 85K chain #status predicted <85K>
F:3945-4545/Domain: 85K chain extracellular #status predicted <EXT>
F:3971-4013/Domain: LDL receptor YWTD-containing repeat homology <YW42>
F:4014-4057/Domain: LDL receptor YWTD-containing repeat homology <YW43>
F:4058-4100/Domain: LDL receptor YWTD-containing repeat homology <YW44>
F:4101-4143/Domain: LDL receptor YWTD-containing repeat homology <YW45>
F:4152-4183/Domain: EGF homology <EG16>
F:4201-4232/Domain: EGF homology <EG17>
F:4237-4268/Domain: EGF homology <EG18>
F:4273-4304/Domain: EGF homology <EG19>
F:4309-4340/Domain: EGF homology <EG20>
F:4345-4375/Domain: EGF homology <EG21>
F:4378-4409/Domain: EGF homology <EG22>
F:4422-4445/Domain: transmembrane #status predicted <TM>
F:4446-4545/Domain: intracellular #status predicted <INT>
F:167,2999/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F:2959/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F:4076,4162,4279/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 24.2%; Score 2113; DB 1; Length 4545;
Best Local Similarity 34.3%; Pred. No. 3.5e-124;
Matches 508; Conservative 265; Mismatches 588; Indels 122; Gaps 51;

Qy 35 LLFANRRDRLVDAGGVKLESTIVVSGLEDAADVPQSGAVYWTVDVSEAIKQTYLNQ 94
Db 1272 IIFSNRHEIRIDLH--KGDYSVLVPLGLRNTIALDFHLSQSALYWTDAVEDKIYRGKLLD 1329

Qy 95 TGA--AVQNVVISGLVSPDGLACDWGKKLYWTDSETRNRIEVANLNGTSRKVLFWQDLQ 152
Db 1330 NGALTSFEWIIQYGLATPEGLAEVIAWAGNIYVWESNLDAQIEVAKLDGTLRTLAGDIEH 1389

Qy 153 PRALADPAHGYMYWTDW-GETPRIERAGMDGSTRKII--VDSDIYWPNGLTIDLERQKL 209
Db 1390 PRALADPRGILFWTDASLPLRIEASMSGARRTHRTGSGGCANGLTVDYLEKRI 1449

Qy 210 YWADAKLSFIHRANLDGSRFQKVVSGS--LTHPFAITLSGDTLYWTDQWTRSIHACNKR 267
Db 1450 LWIDARSDAIYSARYDGSGHMEVLRGHEFLSHPPFAVTLYGGEVYWTWRTWTLAKANKWT 1509

Qy 268 GKKRKEILSALYSPMDQVLVSQRPFPHTRCBEDNGS---CSHLCLLSPSEPPYTCACP 324
Db 1510 GHNVTVQRTNTQFDLQVYHPSQWAPNPCEA--NGGRGPCSHLCLINVRT--VSWACP 1567

Qy 325 TGVOLONGRTCKAGAEVILLARRTLRRISLTDPTDITV-LQVDDIRHAIADYDPL 383
Db 1568 HMKLHKDNTTCYB-FKKFLYAKMEIRGVLDAPYNYIISTVPDINDVTVLDYDAR 1626

Qy 384 EGYVYWTDDVEAIRRAYLDGSGAQTLVNTBNDPDGIADVWARNLYWT--DTGTDRIE 441
Db 1272 IIFSNRHEIRIDLH--KGDYSVLVPLGLRNTIALDFHLSQSALYWTDAVEDKIYRGKLLD 1329

Db 1627 EQRVYMSDVRTQAIKRAFINGTGVETVVSADLPNAHGLAVDWVSRNLFWTSDYTNKKQIN 1686

Qy 442 VTRLINGTSRKTLVSEDEDEPAIALHPVWGLMYWTDNGENPKIECANLNDGQERRVLVNAS 501
Db 1687 VARLDGSGFKNAV--QGLEQPHGLVHPURGLKLYWD--GDN--ISWANNMDSNHTLLFSGQ 1742

Qy 502 LGWPNGLALDQEGKLYWGDAKTDKIEVINVDGTYKRRTL--LEDKLPHFIFGFTLLGDPIY 559
Db 1743 KG-PVGLAIDPESKLYWISSGNTINRCLNDGSELEVIDTMRSQLGKATALAINGDKLW 1801

Qy 560 WTDWQRSIERVHKVKSARDVIIDQLPDL-MGLKAVNVA---KVVGTNPCADRNCGCSHL 615
Db 1802 WADQVSEKMGTCNKADGSGSVLRNSTTLVNMHKYVDESIQLEHEGTNPFCSVNNGDCSQL 1861

Qy 616 CFFTPHATR-CGCPGLELLSDMKTICVPEAFVFTSRAATHRSLEWNN-NDVAIPLTG 673
Db 1862 CLPTSETTRSCMTAGYSLRSGQACGEGVSGFLYSVHEGIRGIPLDNDKSDALVPVSG 1921

Qy 674 VKEASALDFDVSNHIIYWDVSLKTIISRAFNWGSVEHVVEFGLDYPGMAVDMWGNKLY 733
Db 1922 TSLAVGIDFHAENDTIYWDVGLSTISRAKEDQIWRDVTNGIGRVESGIADVDTAGNIY 1981

Qy 734 WADTGNRIEVARLDGQFRQVLVWRDLNPRSLADPTKGYIYWTWGGKPRIVRAFMWG 793
Db 1982 WTDQGFVIEVARLNGSFYVVISQGLDKPRAITVHPKEGYLFTWEMGHYPRIERSRLDG 2041

Qy 794 TNCMTLYD-KVGRANDLTIDYADQRLYKTDLTNMISSNM-LQGER-VWTADDLPHRPF 850
Db 2042 TERVVLVNVSIWPNGISVDYQGGKLYWCARDMDKIERIDLETGENREVLSNNMDMFS 2101

Qy 851 LTOYSDIYIYTDWMLH---SIERADTKTSGRNTLTIQGHLDL-VMDILVFHSSRODGLNDC 906
Db 2102 VSVFEDIYMSD-RTHANGSIKRCCKDNATDSVPLRTGIGVQLKDIKVFNRDRQGTNV 2160

Qy 907 MHNQCGQLCLAIPTGHR-CGCASHYTLDPSSRNCSPPTFTFLFSQKSAISRM-IPDDQ 964
Db 2161 AVANGGQQLCLYRGQGORACACA-HGMLAEDGASCREYAGYLLYSERTILKSHLSDER 2219

Qy 965 HSPDLIILPLHG---LRNVKAIYD-----PLDKFYIYWDGR-QNIKRAKDDGTQPFVL 1013
Db 2220 NLNAPQFDEPHKMKVIALAFDYRAGTSPTGNRIFFSDIHFGNIQOINDGSG---R 2276

Qy 1014 TSLSQGNQPRQPHDLSD--IYSR---TLFWTCEATNTINVHRLSGEAMG-----VV 1061
Db 2277 TTIENVG-----SVEGLAYHGWDTLYWTSYTTSTITRHTVDQTRPGAFERETVIT 2328

Qy 1062 LRGRDRKPRAIYVNAERGYLYFTNMQDRAAKIERAALDGTERTERVLFTTGLIRPVVALVDN 1121
Db 2329 MSGD-DHPRAFVLDCCQNLMTWNNBLHPSIMRAALSGANVLTLIEKDIRTPNGLAIDH 2387

Qy 1122 TLGKLFWVDADLKRIESCDSLGSANRLTLEDANIYVQPLGLTILGKLYWIDRQQQMIEVE 1181
Db 2388 RAELKLYPSDATLDKIERCEYDGSYRVLKSEBPVHPFLAVYGEHIFWTDWVRRAVQRAN 2447

Qy 1182 KTTGD--KRTRIQGRVAHLTGIHAVEEVSLEEFSAHPCARDNGGCSHICAKGDTGRCS 1239
Db 2448 KYVGSDMKLRLVDIPQPMGLIAVANDTNSCLS--FCRINNGGQDLCLLTHQGHVNC 2505

Qy 1240 CPVHLVLLQNLITCGEPTTC-SPDQFACATGEIDCIPGAMRCDFGPECCDDSDSE---G 1294
Db 2506 CRGGRILQEDFTCRVNSSCRAQDEFECANGE--CISFSLTCDGVSHCKDKSDEKPSYCN 2563

Qy 1295 CPVCSAAQFCARGQCYDLRLRCDEADCDQRSDEACD-AICLPNQFRCSGQCVLIKQ 1353
Db 2564 SRRCCKTFRQCNRCRCVSNMLWCVGVDYCGDSDSEIPCNKTACGVGFRCDGSCIGNSS 2623

Qy 1354 QCSFPCIDGSDRLMCEITKPPSDSDSPAHSAGIPVIGITLSLFWMGVYFVQCV--- 1410
Db 2624 RCNOFVDCEDASDEMCATDCSS-----YFLRGVKGVLFPQCERTSLC 2667

Qy 1411 -----VCORYAGAN--GPFPHY-VSGT--PHVPLNFIA-PGG 1442
Db 2668 YAPSWUCD---GANDCGDYSRDERCPGVRPCPLNYFACPSG 2707

1629	Db	MOGTRQIIIVKTKIYWPNTIALDYTTDRVYFADSKLDFIDFVNYDGSGRTOVLASSKFVQ	1688
239	Qy	HPFALTLSGDTLYWTDQWTRSIHACNKRCTGGRKEILSALYS--PMDIQVLSQERQPPF-H	236
1689	Db	HPHALAIPEDMMYYSDRRLQKLQVYPKYPNGTITSEYPSHTFSKALGVVAHPVLQPVIKN	1748
297	Qy	TRCEBONGGSHCLLSPSEBPFYTCAPTGVQLQDNGRTCKAGAEVLLIARTRDILRRIS	356
1749	Db	NPC--STNPCSHLCLLNKNTF--TKPCPMGEKLDASGKKCIDDAKQFLVIIQNTNVFGIE	1805
357	Qy	L-----DTPDFDTIIVLQVDDIRHATAIDYDPLEGVVY-----WTDDEVRA	396
1806	Db	MNSASKEIPVLAGMV-PLISGLGNAPDAAYDALSEMFLEHTNHAKTLAQIITDS--A	1861
397	Qy	IRRAYLDGSGAQTIVNTEI--NDPDGIAVDVARNLTYMTDGTDRIEIVTRLNGTS-RKILV	454
1862	Db	IYRSTVNGGKTKWFSAPVDDAYCYLGFWDNGRNLVVGNKITQITIEIRITQGRQYRSVIL	1921
455	Qy	SEP-----LDEPRAIALHPVMGLMYTWDG---ENPKIECANLDGERRVVLVNASLGWPN	506
1922	Db	SNDQSPAVVTPVAIYAVDADKGVYFVLDRGGGAADAKVARAGLDGSLPLVIASNDLAE	1981
507	Qy	GLALDIQEGKLYMGDAKTKIEVINVDGTRKRTLLEDKPHIFEGTLGDFIYWTQWQR	566
1982	Db	HIAIDTNTVRFPSEAKAGRISSVYDQDRHYVLSDGGRQPNGLAFYGBRLFIASAFD	2041
567	Qy	SIERVHKVKA-----SRDVIIDQLPDLMLKAVNVKVTGVPNCADRNGGCSHLC	616
2042	Db	SIE-VATINGDSOPPQWTHFKDQV-----ENLANIKALQPRASSSGHPCHINNGNDHIC	2095
617	Qy	FPTPHATR--CGCPIGLELLSDMKTCIVPRAFLVFTSRAAIHRISL--ETNNNDVAIPLTVG	674
2096	Db	IPLMPAQRCTCTCANGVYVDCGTCKFLFDESFIIVATKVIQYIPIDETOSKGVAMEPIGG	2155
675	Qy	KEASALDFDVSNHHIYWTDS--LKTISRAPWNGSSVEHV--EFGLDYPEGMADVMMGX	730
2156	Db	LSITGVDDYDYESKTIIVBASGINKGITAYTIGESSPRAVIRDSISGLTIKSIAIDWNY	2215
731	Qy	NLYWA--DTGTRNIEVARLDGPOQVLVWRDLNPRSLALDPTPKYIYTEMGCKPRIVR	788
2216	Db	NMYFINHAERTNIEVSKLDGTVRKILLTITKTETPSSIAVDPVSRLYWADQCKPTQIR	2275
789	Qy	AFMDGNCMTLVDK--VGRANDLTIDVADQRLYWTDLDTNMIESSNMLGQERVVIADLPH	847
2276	Db	SFLDGSRRREVIVSSGIAEPTDLVDVAVASKMIYWSDAKMDGIYVRSTGGTPELVRSDIAS	2335
848	Qy	PPELQYSDIYIYTDWNTHASIERADKTSGRN-----TLQGHLDVMDILVFEHSRQ	900
2336	Db	AGVALHQNMVYTDNRLKLFRA--TSKPNQTSLLSPPTTVAASLUKDIDGVAVFSSNNQ	2393
901	Qy	-----DGLNDCMHNNGOQCLAIIPGGH--RCGCASHYTLDPSSRNCSPPTTFL	948
2394	Db	PRASSPCQITDNL-----RKSPCTQLCFATPGTQTCTSCARGVL---KGRTCEEPDTYI	2445
949	Qy	LFSQKSAISRM--PDQHSPLDILPLHGLURNVKAIDYDPLDKFIYVDGQRNIKRAQD	1006
2446	Db	MFSGDGKIIDVATEPDVKASRLPKDPFPEISNLQTFDVDVNLRRVYFV-----VESPVGV	2500
1007	Qy	GTQPFVLTLSQGNP-----DROPH-----DLSIDIYRSLFTWCEATNTINVHR	1052
2501	Db	NISWF---SMNNAENPRLVFGASKQPHAKEIRHISDMKLDMLTKQIYFTTGRGCKVMAID	2557
1053	Qy	LSGEAMGVVLGRDRKPRAIVVNAERGILYFTNMQDRAA-----KIERAALDGTREVL	1106
2558	Db	TAGEHLSTTASG--DWTYALADPCPSGLLFWSDSGYKTSGLLYEPRIERNLAGSGRKVI	2615
1107	Qy	FTTGLIRPVALVVDNTLKGFLFWVDADLKRIESCDLSGANRLTLIEDANIVOPLGLTILGKH	1166
2616	Db	VSESIISLPAALAVDFRNQKIYADVNRLNIEVADYDQGNRKVI--ASGVRAKSLDIWDWR	2673
1167	Qy	LYWIDRQOQMIERVEKTTGD-----KRTRIQRV---AHLTGIHAVEEVSLSEFSAHPC	1217
2674	Db	LYMSDPLSNGVFRIDKESGSLGVNYSDBRIIPGALRVFASIESDVRTRNQV--CNALTSQLC	2732

QY 1218 ARDNGCSHICAKGD-----GTRCSC-----PVLHVLQN----- 1249
Db 2733 KTDNGGCDLCTVAVDDIGLAASKVQCSCNDTYELVBFQKDYPTQCVLRGNSBPAKEC 2792
QY 1250 -----LL--TCGBPPCTCS-----PDQFACATGBIDCIPGAW 1278
Db 2793 LPPYNFQCGDSCILLGATCDKSKPCADASDENPNYCNTRSCPDYNLCTNR-RCIDSAK 2851
QY 1279 RCDGFPECDQSDREGCP---VCSAAQFPARGQCVRDLRLRCDGBADCCDR--SDEA--- 1330
Db 2852 KCHNIDDCGDSDELDCPSAVACABGTFCPSNGHCINQTKVCDGHNDCHDEQVSDSLAT 2911
QY 1331 -----DCDAI-----CLPNQPRCASG 1346
Db 2912 CGPLPIDCRGVKVRCPNNICIQPADLDCGYDDCGKADENQLFCMNQCCAQHYVRCPSG 2971
QY 1347 QCVLKIQCCDSPDCIDGSDLMCEITKPPSDSPAHSSAIGPVGIGIILSLFVMGVVFPV 1406
Db 2972 RCIPETWQCDGNDSCDGDWDETHNCT-----DTAGKIKCVGDYL-----FQ 3013
QY 1407 QORVVC 1412
Db 3014 CDNLAK 3019

RESULT 10

T46336

Hypothetical protein DKFp43400213.1 - human

C:Species: Homo sapiens (man)

C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004

C:Accession: T46336

R:Koehret, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23037

A:Accession: T46336

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-252 <AA>

A:CROSS-references: UNIPROT:O9NSY4; EMBL:AL137651

A:Experimental source: adult testis; clone DKFp43400213

C:Genetics:

A:Note: DKFp43400213.1

C:Superfamily: low density lipoprotein receptor-related protein; EGF homology; LDL recep

Query Match 11.7%; Score 1024; DB 2; Length 252;
Best Local Similarity 83.5%; Pred. No. 2.8e-57;
Matches 193; Conservative 16; Mismatches 22; Indels 0; Gaps 0;

QY 383 LEGVYVWTDDEVRAIRRAYLDGSGAOTLVNTEINDPGIADVWVARNLYTDTGTDRTEV 442
Db 1 MEGHYVWTDDEWAIIRRAYLDGSGAOTLVNTEINDPGIADVWVARNLYTDTGTDRTEV 60
QY 443 TRLNGTSKILVSEDLDEPRAIALHPVGMGLMYWTWNGENPKIECANLDGQRRVNVNASL 502
Db 61 TCLNSTSHKILVSEDMDEPRAIALHPENGLTYWIDWNGENPEIKRANLDQRLVNVNASL 120
QY 503 GWPNGALDLOEGKLYWGDAKTDKIEVINVDGTRKRTLEKLEPHIFGFTLLGDFIYWD 562
Db 121 GWPNGALDLOEGKLYWGDAKTDKIEAISVDTRKQTLKDLKLEPHIFRFTLLGDFIYWA 180
QY 563 WQRSIERVHVKASRDVIIIDQLPDLMLKAVNVKVVGTNTPCADRNGGCS 613
Db 181 WQHSIEKRVHVKANRDVIIIDQLPDLMLKAVNVKVVGTNTPCADRNGGAA 231

RESULT 11

T18308

probable vitellogenin receptor - yellow fever mosquito

C:Species: Aedes aegypti (yellow fever mosquito)

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T18308

R:Sappington, T.W.; Kokoza, V.A.; Cho, W.L.; Raikhel, A.S.

Proc. Natl. Acad. Sci. U.S.A. 93, 8934-8939, 1996
A>Title: Molecular characterization of the mosquito vitellogenin receptor reveals unexpect
A:Reference number: Z18874; MUID:96392343; PMID:8799131
A:Accession: T18308
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1847 <SNP>
A:CROSS-references: UNIPROT:O76952; EMBL:L77800; NID:g3386582; PID:g3386583; PIDN:AAC2845

C:Function:

A>Description: involved in endocytosis
F:45-79/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:84-120/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:132-167/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:174-207/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:211-249/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F:256-287/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:298-329/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F:332-363/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F:367-398/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:402-433/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F:437-468/Domain: LDL receptor ligand-binding repeat homology <LDL11>
F:472-503/Domain: LDL receptor ligand-binding repeat homology <LDL12>
F:507-538/Domain: LDL receptor ligand-binding repeat homology <LDL13>

Query Match 10.5%; Score 914.5; DB 2; Length 1847;

Best Local Similarity 24.9%; Pred. No. 4.1e-49;

Matches 299; Conservative 150; Mismatches 386; Indels 365; Gaps 46;

QY 299 CEEDN-----GGCSHLCLLSPSEPFYTCACPTGVQLQONGRTCKA--GAEEVLLIARRTD 351
Db 287 CEDINECEYGLCSQGCENTPGS--FKCTCVDFKFKDDSRTELDSDSTELLVYTTQKS 344
QY 352 LRRISLTPPTDVLQVDDIRHAIADYDPLEGYVYVWD--DEVRAIRRAYLDGSGAQT 409
Db 345 IGGILHLNTHQYVY--AKDLSQVIGSYD--GRHVYVWTDISPKTESLSDGTREL 399
QY 410 LVNTEINDPGIADVWVARNLYTDTGTDRTEVRLNGTSKILVSEDLDEPRAIALHPV 469
Db 400 LITSGLASPEDELDWLTGNIYFSDSGHMMIAVCSNNGVHCTILIQTLHPRGIALMPQ 459
QY 470 MGLMYWTDWGENPKIECANLDGQRRVNVNASLGNLWPNGLALDLOEGKLYWGDAKTDKIEV 529
Db 460 NGTLFYSWGDGNDAMIGAAGDGKRNKILIEQDIHWPNGCLDWPNGRLYVWDKXKIES 519
QY 530 INVDTGTRTLEKLEPHIFGFTLLGDFIYWDQWRSIERVHVKA-SRDVIID--QLP 586
Db 520 IRVDGTRNVTLADVLKHPFSIAVFNDRLYWSDWDTKSIQSCDKPFGKDKRIVVHQRIF 579
QY 587 DL-----MGLKAVNVKVVGTNTPCADRNGGCSHLCTFFTPH-ATRCGCPIGLELLSDMKTG 640
Db 580 DVHIYHSGLP-----KGDPCL--GTFCSHLCLLAPNDSYSCACPYGSLKADKHSR 631
QY 641 -IVEAPFLVFTSRAAHRIISLENNNDVAIPLTGKESALD-----FVSNHHI 689
Db 632 ETVKRQYLLV--GIANYLVLTETQ-----FGHESSQADAYQIFFRMAFNSITGEI 682
QY 690 YWTDVSLKTSIRAPMNGSSVHVVEFGLDYPEGNAVDMGKILYWDGTNRIEVARLDG 749
Db 683 FVADNRQAKFTVDPKTKSSQKLITTTGIGNISALAFDLGNLNYTDSERSTVEFSLQT 742
QY 750 QFROVLY-WRDLONPRSIALDPTKGYIY-----WTEWGGKPRIVRAF 790
Db 743 RHRAILQHYLGQDIPVGLAIVSEMGKFIARLSPVPVPHTHIDRLDWTGRGPHLVIEER 802
QY 791 MDGTNCMTLVKGRANDLTDYADQRLYWDLTDTNMIESNNMLGQRRVVIADLPHFPF 850
Db 803 LSGNGSFNV-----IDRLRTVYVWDMGSSKIEFTSYEGDTRHLFRFLRPLVS 852
QY 851 LTQVSDVIYVWDMNLSHIERADKTS--GRNRTLQGHLD-----FVMDILVHFS--SRQ 900
Db 853 IAVGDSIFWCTVSKKLYWSDKHLNLTGVTTKIT---IDKPPYGAFFDEIVLLGSGQPLQRY 909
QY 901 DGLNDCMHNNGQCCQLCLAIPEGHRCGASHYTLDPSSRNCSPPTTFLFSQKSAISRM 960


```

Db      910 D--HPCMKQGGCCHIC--VPAG-----MYSSACICPTGMIFFSPK----- 946
Qy      961 PDDQHSPLILPLHLNRVAIDYDPLDKFYIYWDGRQNIKRAKDDGTQFVLTSLSQGG 1020
Db      947 -----NTTCID-----
Qy      1021 NPDRQPHLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVVLR--GDRKPRAIIVNAER 1078
Db      953 -----AID-----CEPKCT-----SGECLTISKRCNGKD----- 977
Qy      1079 GYLFTNMQDRAAKIERAALDGTREVERLFTTGLIRPVVALVVDNTLGLFWVDADLKRIES 1138
Db      978 -----CADGSD-----
Qy      1139 CDLSGANRLTLEDANIVQPLGLTLGKHLVY-----IDRQOMIERVEKTTGDK 1187
Db      987 CDEAG-----QPKQL-----HCQYDEFMCADKSKCIDQTRCDEHVDGSD 1029
Qy      1188 RTRIQRVAHLTGTHAVEEVSLEBFSAPCARDNGGC---SHICIAKGDGTPRC---SCP 1241
Db      1030 EMKEG-YDRGTGCH-----BHQACPGMCIIDVNTLC-----DGFPCDCLGSD 1073
Qy      1242 VHLVLLQNLTLTCEPPTCSPDPQACATGEIDCIPGAWRCDFPCCDQSDSDE-EGCPV--- 1297
Db      1074 VGCTDLTN--EKSNAITCGPLMPCNMGQ--CIPKWECDGNPCTDGSDEHDKLTKTD 1129
Qy      1298 CSAAQFPCARGQCVDLRLRCDEADQDRSDEACD-----
Db      1130 CGAGFTKALGHCHIEDRLCLDGNNDGNSDELNCKVELEPCVGLGEDDNTKYLCPRS GK 1189
Qy      1334 -----AICLPNPRCASGCVLKKQCCQDFPCCDGSDELNC 1370
Db      1190 CLDIARVCRGTAECPGDEDEAGSCNGLQBFQCKSGKIRKEWRCEKVDCCDGSDEVC 1249

RESULT 12
T13171
probable vitellogenin receptor - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C;Accession: T13171
R;Schonbaum, C.P.; Lee, S.; Mahowald, A.P.
Proc. Natl. Acad. Sci. U.S.A. 92, 1485-1489, 1995
A;Title: The Drosophila yolkless gene encodes a vitellogenin receptor belonging to the 1
A;Reference number: 217627; MUID:95183490; PMID:7878005
A;Accession: T13171
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1984 <SCH>
A;Cross-references: UNIPROT:P98163; EMBL:U13637; NID:G535345; PID:G535346; PIDN:AAB60217
C;Genetics:
A;Gene: yl
A;Cross-references: FlyBase:FBgn0004649
A;Map position: 1
F;90-124/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F;129-166/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F;184-220/Domain: LDL receptor ligand-binding repeat homology <LD10>
F;227-262/Domain: LDL receptor ligand-binding repeat homology <LD11>
F;266-304/Domain: LDL receptor ligand-binding repeat homology <LD12>
F;1025-1062/Domain: LDL receptor ligand-binding repeat homology <LD11>
F;1074-1109/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F;1118-1152/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F;1158-1193/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F;1198-1232/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F;1243-1279/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F;1283-1318/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F;1340-1375/Domain: LDL receptor ligand-binding repeat homology <LDL9>

Query Match      10.5%; Score 913.5; DB 2; Length 1984;
Best Local Similarity 23.8%; Pred. No. 5.3e-49;
Matches 305; Conservative 162; Mismatches 434; Indels 381; Gaps 47;
```

```

Qy      299 CEEDNGGCSHCLLSPSEPPYTCACPTGVOLQNGRTCKA---GABE---VLLLRRTDL 352
Db      352 CKQDDLLCSQGC--ENTSGGYRCVCDAGYLLDKDNRICRAVYVYSGKEQQPLLLLTTQWTI 409
Qy      353 RRLSLDPTDIDVLQV--DDIRHAIAIDYDPLEGYVYWD--DEVRAIRRAYLDSGGAQT 409
Db      410 MGHRLRDNVRNHYVQVAGNLSKVIGVAYD--GSHIYWTNIQNEAESIVKANGDGSNAEI 467
Qy      410 LVNTEINDPGIADVWVARNLYWTDCTDRTEVTRNGTSRKILLVSDLDLPPRAIALHPV 469
Db      468 LLTSGLDAPEDLADVLTQNIYFSDNIMRHIAVCSNDGLNCAVLTVQDVHQPRLSAWVPQ 527
Qy      470 MGLMYWTDWGENPKIECANLDGQERRVLVNASLGPNGLALDLOEGKLYWGDADTKTKEV 529
Db      528 KGLMFWTDWGEKPMIGRASMDGSRSPVSDNIEMPNGLALDMEHQRIYWDVAKLGSVQT 587
Qy      530 INVDTGRTLLDKLPHIFGFTLLGDFIYWTDMQRHSIERVHVKKVAKASRDVIIDQLPDL 589
Db      588 VRPDGTGRRTVLQGLMKAHPYGLAIFEDQLYMSDWATKSVHACHKFSQKDHRIILAKDR 647
Qy      590 GLKAVNVKVVGTNPACDRNGGCSHLCFFTP-----HATRCGCPGLELLSMDMTKCI 644
Db      648 AVHIYHPAK--OPNSPHGECENATCSHCLLAEPEIGGHS--CACPDGMLAPDHRRCMLME 704
Qy      645 ---AFLVFTSRAAIHRISLETNNNDVAIPLTVGKEAS-----ALDPDVSNNNHYWTDV 694
Db      705 KRQLFI-----GLGQVLLIEHT-----AFGAHQVSKSYTLPCLINEVYNRINGS 753
Qy      695 SLKTIISRAFWNGSSVEH-----VVEFGLDYPEGMAVDMGNKLYWADTGTNRRIEVAR 750
Db      754 IADNDQRLILEFQEPESHESVLRSLNGLNSALAFDHLRNLRYWADTERAVIEVLSLQTR 813
Qy      751 FRQVL--VWRDLNPRSLALDPTKGIYWT-----EWGKPRIVPAFMD--GTN 795
Db      814 HRALRFPFGQEPVIGLTVMPAEGYLVYLAKKRHSIDKIPISGKGEQVHFDDLDGDD 873
Qy      796 CMTILVDKVGGRANDLITIDYADQRLYWTDLDTNMISSNNMLGQERVIADDLPHFPGLTO 855
Db      874 DIKLV-----TDYETQTFWSDSLGRISYVYVPHSOIFRGKLRPRPSLAMVH 923
Qy      856 DYYIWTDMNLHSIERADKTSGRNRTLQGHLDLFWD-----ILVFH 896
Db      924 HDLFWNELGTPIRYWTHKSNMGRPKVID-----IMEKDDPAAIMPYPVPATPNGIPLA 978
Qy      897 SSRQDGLNDCHWNGQCGQLAIPGGRH--CGCASHYTL--DPSSRNCSPPTTFLFSQKS 954
Db      979 SPVQESHPCQQNGGCSHCICVGGPYHSLCPAGFVYRDAGNRTC----- 1025
Qy      955 AISRMIPDDQ--HSPDLILPLH---GLRNVKAIDYDPLDKFYIYWDGRQNIKRAKDDGTQ 1010
Db      1026 -VEALDCEFRCHSGECLTMNRCNGRDC-----VDNSDEMNCDEHRHKP 1070
Qy      1011 FVLTSLSQGNQPRQPHDLSDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPR 1070
Db      1071 KVLCSPSQ-----FACH-----SSEQ----- 1086
Qy      1071 AIVVNAERGLYFTNMQDRAAKIERAALDGTREVERLFTTGLIRPVVALVVDNTLGLFW 1130
Db      1087 --CVDKER----- 1092
Qy      1131 ADLKRIESCDSGANRLTLEDANIVQPLGLTLGKHLWIDRQOQMTERVEKTTGDKRTR 1190
Db      1093 -----RCD---NRKCDHDS-----DQHCKEFKDS----- 1115
Qy      1191 IQGRVAHLTGTHAVEEVSLEBFSAPCARDNGGC---SHICIAKGDGTPRCSPVHLVLL 1247
Db      1116 ---KKCH---VH-----OHGC--DNGKCVDSLSVC-----DGTNDGCDN----- 1146
Qy      1248 QNLLTCGEPTCSPDPQACATGEIDCIPGAWRCDFPCCDQSDSDEGCPV---CSAAQFP 1304
Db      1147 SDELLCEATLRCFPMFQCGSG--SCIAGSWECDGRIDCSGSDGDEHDKCVHRSPDPMHR 1204
Qy      1305 CARQCVDRLRLRCDEADQDRSDEADCD-----ICLPN----- 1339
```


Qy 441 EVTRNGTSRKILVSELDDEPRATIALHPVGMVLMYWTDMGENPKTECANLGDQERRVLNA 500
Db 580 GRSDNGRSLIKTIKENISQPRGTAHVHAKRFLWTDIGINPRIESSLQGLGRUVIASS 639
Qy 501 SLGPNGLALDQEGKLYWGDAKTDKIEVINVDGTRKRTLLDKLPHIFGFTLLGDFIYW 560
Db 640 DLWPSGITIDFLTDKLYWCAKQSVIEWANLGDGSKRRRLTQNDVGHFFAVAVPEDIYWF 699
Qy 561 TDWQRSTIRVHKVVASRDVIDQLPDLM----GLKAVNAKVGTNPNCADRNCGCSHLG 616
Db 700 SDWAMPISIRVKN-RTGDRV--RLQGSMLKPSLLVVVHPLAKPGADPCLYQNGGCEHC 756
Qy 617 FFTPHATRCGPIGLELLSDMKTCI 641
Db 757 KRLGTAWCSREGPMKASDKTKCL 781

RESULT 15
EGMSMG
epidermal growth factor precursor - mouse
N:Alternate names: urogastone precursor
C:Species: Mus musculus (house mouse)
C:Date: 30-Nov-1980 #sequence revision 11-Aug-1983 #text_change 09-Jul-2004
R:Accession: A94272; A93304; A92118; A01387
R:Scott, J.; Urdea, M.; Quiroga, M.; Sanchez-Pescador, R.; Fong, N.; Selby, M.; Rutter, Science 221, 236-240, 1983
A>Title: Structure of a mouse submaxillary messenger RNA encoding epidermal growth factor
A:Reference number: A94272; MUID:83223630; PMID:6602382
A:Accession: A94272
A:Molecule type: mRNA
A:Residues: 1-1217 <SCO>
A:Cross-references: UNIPROT:P01132; GB:J00380; NID:g192993; PIDN:AAA37539.1; PID:g309210
R:Gray, A.; Dull, T.; Ullrich, A.
Nature 303, 722-725, 1983
A>Title: Nucleotide sequence of epidermal growth factor cDNA predicts a 128,000-molecular weight protein
A:Reference number: A93304; MUID:83219309; PMID:6304537
A:Accession: A93304
A:Molecule type: mRNA
A:Residues: 1-789, 'Y', 791-1047, 'S', 1049-1168 <GRA>
A:Cross-references: GB:J00380
A>Note: The sequence shown by these authors differs from residues 1134-1168 due to an inference of Scott et al.)
R:Savage Jr., C.R.; Inagami, T.; Cohen, S.
J. Biol. Chem. 247, 7612-7621, 1972
A>Title: The primary structure of epidermal growth factor.
A:Reference number: A92118; MUID:73048516; PMID:4636327
A:Accession: A92118
A:Molecule type: protein
A:Residues: 977-1029 <SAV>
A>Note: residues 1024-1029 are not required for full biological activity in vivo
R:Savage Jr., C.R.; Hash, J.H.; Cohen, S.
J. Biol. Chem. 248, 7669-7672, 1973
A>Title: Epidermal growth factor. Location of disulfide bonds.
A:Reference number: A92144; MUID:74025498; PMID:4750422
A:Contents: annotation; disulfide bonds
A:Comment: Epidermal growth factor (EGF) stimulates the proliferation and differentiation of epithelial cells.
A:Comment: The EGF precursor is found in kidney as a receptor-like membrane-bound protein.
A:Comment: The active growth factor from this submaxillary gland protein stimulates the proliferation and differentiation of epithelial cells.
A:Superfamily: epidermal growth factor precursor; EGF homology; LDL receptor WYTD-containing repeat; duplication; growth factor; tandem repeat; transmembrane protein
C:Keywords: duplication; growth factor; tandem repeat; transmembrane protein
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-1217/Product: epidermal growth factor precursor, membrane-bound form #status predicted <EXT>
F:29-1038/Domain: extracellular #status predicted <EXT>
F:50-485/Region: EGF precursor long repeat
F:53-92/Domain: LDL receptor WYTD-containing repeat homology <YW01>
F:93-134/Domain: LDL receptor WYTD-containing repeat homology <YW02>
F:135-176/Domain: LDL receptor WYTD-containing repeat homology <YW03>
F:177-217/Domain: LDL receptor WYTD-containing repeat homology <YW04>
F:218-262/Domain: LDL receptor WYTD-containing repeat homology <YW05>
F:263-307/Domain: LDL receptor WYTD-containing repeat homology <YW06>

F:324-360/Domain: EGF homology #status atypical <EG1>
F:366-401/Domain: EGF homology <EG2>
F:407-442/Domain: EGF homology <EG3>
F:445-482/Domain: EGF homology <EG4>
F:486-961/Region: EGF precursor long repeat
F:489-529/Domain: LDL receptor WYTD-containing repeat homology <YW07>
F:530-572/Domain: LDL receptor WYTD-containing repeat homology <YW08>
F:573-615/Domain: LDL receptor WYTD-containing repeat homology <YW09>
F:616-659/Domain: LDL receptor WYTD-containing repeat homology <YW10>
F:660-700/Domain: LDL receptor WYTD-containing repeat homology <YW11>
F:701-743/Domain: LDL receptor WYTD-containing repeat homology <YW12>
F:751-786/Domain: EGF homology <EG5>
F:842-875/Domain: EGF homology <EG6>
F:881-917/Domain: EGF homology <EG7>
F:923-958/Domain: EGF homology <EG8>
F:977-1029/Product: epidermal growth factor #status experimental <EGF>
F:982-1018/Domain: EGF homology <EG9>
F:1039-1063/Domain: transmembrane #status predicted <TMM>
F:1064-1217/Domain: intracellular #status predicted <INT>
F:347-360,366-377,373-386,388-401,407-418,414-427,429-442,445-457,453-467,469-482,751-761,763-777,781-791,793-805,807-819,821-833,835-847,849-861,863-875,877-889,891-903,905-917,919-931,933-945,947-959,961-973,975-987,989-1001,1003-1015,1017-1029,1031-1043,1045-1057,1059-1071,1073-1085,1087-1100,1102-1114,1116-1128,1130-1142,1144-1156,1158-1170,1172-1184,1186-1198,1199-1211,1213-1225,1227-1239,1241-1253,1255-1267,1269-1281,1283-1295,1297-1309,1311-1323,1325-1337,1339-1351,1353-1365,1367-1379,1381-1393,1395-1407,1409-1421,1423-1435,1437-1449,1451-1463,1465-1477,1479-1491,1493-1505,1507-1519,1521-1533,1535-1547,1549-1561,1563-1575,1577-1589,1591-1603,1605-1617,1619-1631,1633-1645,1647-1659,1661-1673,1675-1687,1689-1701,1703-1715,1717-1729,1731-1743,1745-1757,1759-1771,1773-1785,1787-1799,1801-1813,1815-1827,1829-1841,1843-1855,1857-1869,1871-1883,1885-1897,1899-1911,1913-1925,1927-1939,1941-1953,1955-1967,1969-1981,1983-1995,1997-2009,2011-2023,2025-2037,2039-2051,2053-2065,2067-2079,2081-2093,2095-2107,2109-2121,2123-2135,2137-2149,2151-2163,2165-2177,2179-2191,2193-2205,2207-2219,2221-2233,2235-2247,2249-2261,2263-2275,2277-2289,2291-2303,2305-2317,2319-2331,2333-2345,2347-2359,2361-2373,2375-2387,2389-2401,2403-2415,2417-2429,2431-2443,2445-2457,2459-2471,2473-2485,2487-2499,2501-2513,2515-2527,2529-2541,2543-2555,2557-2569,2571-2583,2585-2597,2599-2611,2613-2625,2627-2639,2641-2653,2655-2667,2669-2681,2683-2695,2697-2709,2711-2723,2725-2737,2739-2751,2753-2765,2767-2779,2781-2793,2795-2807,2809-2821,2823-2835,2837-2849,2851-2863,2865-2877,2879-2891,2893-2905,2907-2919,2921-2933,2935-2947,2949-2961,2963-2975,2977-2989,2991-3003,3005-3017,3019-3031,3033-3045,3047-3059,3061-3073,3075-3087,3089-3101,3103-3115,3117-3129,3131-3143,3145-3157,3159-3171,3173-3185,3187-3199,3201-3213,3215-3227,3229-3241,3243-3255,3257-3269,3271-3283,3285-3297,3299-3311,3313-3325,3327-3339,3341-3353,3355-3367,3369-3381,3383-3395,3397-3409,3411-3423,3425-3437,3439-3451,3453-3465,3467-3479,3481-3493,3495-3507,3509-3521,3523-3535,3537-3549,3551-3563,3565-3577,3579-3591,3593-3605,3607-3619,3621-3633,3635-3647,3649-3661,3663-3675,3677-3689,3691-3703,3705-3717,3719-3731,3733-3745,3747-3759,3761-3773,3775-3787,3789-3801,3803-3815,3817-3829,3831-3843,3845-3857,3859-3871,3873-3885,3887-3899,3901-3913,3915-3927,3929-3941,3943-3955,3957-3969,3971-3983,3985-3997,3999-4011,4013-4025,4027-4039,4041-4053,4055-4067,4069-4081,4083-4095,4097-4109,4111-4123,4125-4137,4139-4151,4153-4165,4167-4179,4181-4193,4195-4207,4209-4221,4223-4235,4237-4249,4251-4263,4265-4277,4279-4291,4293-4305,4307-4319,4321-4333,4335-4347,4349-4361,4363-4375,4377-4389,4391-4403,4405-4417,4419-4431,4433-4445,4447-4459,4461-4473,4475-4487,4489-4501,4503-4515,4517-4529,4531-4543,4545-4557,4559-4571,4573-4585,4587-4599,4601-4613,4615-4627,4629-4641,4643-4655,4657-4669,4671-4683,4685-4697,4699-4711,4713-4725,4727-4739,4741-4753,4755-4767,4769-4781,4783-4795,4797-4809,4811-4823,4825-4837,4839-4851,4853-4865,4867-4879,4881-4893,4895-4907,4909-4921,4923-4935,4937-4949,4951-4963,4965-4977,4979-4991,4993-5005,5007-5019,5021-5033,5035-5047,5049-5061,5063-5075,5077-5089,5091-5103,5105-5117,5119-5131,5133-5145,5147-5159,5161-5173,5175-5187,5189-5201,5203-5215,5217-5229,5231-5243,5245-5257,5259-5271,5273-5285,5287-5299,5301-5313,5315-5327,5329-5341,5343-5355,5357-5369,5371-5383,5385-5397,5399-5411,5413-5425,5427-5439,5441-5453,5455-5467,5469-5481,5483-5495,5497-5509,5511-5523,5525-5537,5539-5551,5553-5565,5567-5579,5581-5593,5595-5607,5609-5621,5623-5635,5637-5649,5651-5663,5665-5677,5679-5691,5693-5705,5707-5719,5721-5733,5735-5747,5749-5761,5763-5775,5777-5789,5791-5803,5805-5817,5819-5831,5833-5845,5847-5859,5861-5873,5875-5887,5889-5901,5903-5915,5917-5929,5931-5943,5945-5957,5959-5971,5973-5985,5987-5999,6001-6013,6015-6027,6029-6041,6043-6055,6057-6069,6071-6083,6085-6097,6099-6111,6113-6125,6127-6139,6141-6153,6155-6167,6169-6181,6183-6195,6197-6209,6211-6223,6225-6237,6239-6251,6253-6265,6267-6279,6281-6293,6295-6307,6309-6321,6323-6335,6337-6349,6351-6363,6365-6377,6379-6391,6393-6405,6407-6419,6421-6433,6435-6447,6449-6461,6463-6475,6477-6489,6491-6503,6505-6517,6519-6531,6533-6545,6547-6559,6561-6573,6575-6587,6589-6601,6603-6615,6617-6629,6631-6643,6645-6657,6659-6671,6673-6685,6687-6699,6701-6713,6715-6727,6729-6741,6743-6755,6757-6769,6771-6783,6785-6797,6799-6811,6813-6825,6827-6839,6841-6853,6855-6867,6869-6881,6883-6895,6897-6909,6911-6923,6925-6937,6939-6951,6953-6965,6967-6979,6981-6993,6995-7007,7009-7021,7023-7035,7037-7049,7051-7063,7065-7077,7079-7091,7093-7105,7107-7119,7121-7133,7135-7147,7149-7161,7163-7175,7177-7189,7191-7203,7205-7217,7219-7231,7233-7245,7247-7259,7261-7273,7275-7287,7289-7301,7303-7315,7317-7329,7331-7343,7345-7357,7359-7371,7373-7385,7387-7399,7401-7413,7415-7427,7429-7441,7443-7455,7457-7469,7471-7483,7485-7497,7499-7511,7513-7525,7527-7539,7541-7553,7555-7567,7569-7581,7583-7595,7597-7609,7611-7623,7625-7637,7639-7651,7653-7665,7667-7679,7681-7693,7695-7707,7709-7721,7723-7735,7737-7749,7751-7763,7765-7777,7779-7791,7793-7805,7807-7819,7821-7833,7835-7847,7849-7861,7863-7875,7877-7889,7891-7903,7905-7917,7919-7931,7933-7945,7947-7959,7961-7973,7975-7987,7989-8001,8003-8015,8017-8029,8031-8043,8045-8057,8059-8071,8073-8085,8087-8099,8101-8113,8115-8127,8129-8141,8143-8155,8157-8169,8171-8183,8185-8197,8199-8211,8213-8225,8227-8239,8241-8253,8255-8267,8269-8281,8283-8295,8297-8309,8311-8323,8325-8337,8339-8351,8353-8365,8367-8379,8381-8393,8395-8407,8409-8421,8423-8435,8437-8449,8451-8463,8465-8477,8479-8491,8493-8505,8507-8519,8521-8533,8535-8547,8549-8561,8563-8575,8577-8589,8591-8603,8605-8617,8619-8631,8633-8645,8647-8659,8661-8673,8675-8687,8689-8701,8703-8715,8717-8729,8731-8743,8745-8757,8759-8771,8773-8785,8787-8799,8801-8813,8815-8827,8829-8841,8843-8855,8857-8869,8871-8883,8885-8897,8899-8911,8913-8925,8927-8939,8941-8953,8955-8967,8969-8981,8983-8995,8997-9009,9011-9023,9025-9037,9039-9051,9053-9065,9067-9079,9081-9093,9095-9107,9109-9121,9123-9135,9137-9149,9151-9163,9165-9177,9179-9191,9193-9205,9207-9219,9221-9233,9235-9247,9249-9261,9263-9275,9277-9289,9291-9303,9305-9317,9319-9331,9333-9345,9347-9359,9361-9373,9375-9387,9389-9401,9403-9415,9417-9429,9431-9443,9445-9457,9459-9471,9473-9485,9487-9499,9501-9513,9515-9527,9529-9541,9543-9555,9557-9569,9571-9583,9585-9597,9599-9611,9613-9625,9627-9639,9641-9653,9655-9667,9669-9681,9683-9695,9697-9709,9711-9723,9725-9737,9739-9751,9753-9765,9767-9779,9781-9793,9795-9807,9809-9821,9823-9835,9837-9849,9851-9863,9865-9877,9879-9891,9893-9905,9907-9919,9921-9933,9935-9947,9949-9961,9963-9975,9977-9989,9991-10000

Db	643	ASSNLEPSGITIDYLTDTLYWCDTKRSVIEANLDGSKRRRLIQNDVGHPPFLAVFEDH	702
Qy	558	IYWTQWQRSTERYHK-----VKASRDVIIQDLPDLMLGLKAVNVAKVGTNP	604
Db	703	LWVSDWAIPSVIRVNKRTGQNRVRLQGSMLPSSLVVVHPL-----AK-PGADP	750
Qy	605	CADRRGGCSHLCFPTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAA--IHRISLET	662
Db	751	CLYRNGGCEHICQESLGTARCLCREGFVKAWDGKML-PQDYPILSGENADLSKEVTSL	809
Qy	663	NNNDVAIPLTGVKEASALDPDVSNHHIYWTDVSLKTSRAFNGSSVEHVVEFGLDYPEG	722
Db	810	NSTQAEVPDDDGTSSTL-----VAEIMVSGMYEEDD	841
Qy	723	MAVDMGKNLYWADTGTNRIEVARLDGQFRQ-----VLVWRDLNPRSLALDPTK	772
Db	842	CGPGCGSHARCVSOG-ETAEQCCLKGFARDGNLCSIDIECVLARSDCPSTSSRCINTEG	900
Qy	773	GYIYWTWGGKPRIVRAFMDGTNCMTL	799
Db	901	GYVCRCSGEGEG-----DGISCFDI	920

Search completed: February 17, 2005, 01:35:57 .
Job time : 69 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2005, 01:35:01 ; Search time 180 Seconds
(without alignments)
2931.666 Million cell updates/sec

Title: US-09-931-375A-2

Perfect score: 8740

Sequence: 1 MEAAPGPPWPLLLLLLLLLL.....TERTSYFHLFPPTDSS 1615

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1376875 seqs, 326749119 residues

Total number of hits satisfying chosen parameters: 1376875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8740	100.0	1615	10	US-09-931-375A-2
2	8736	100.0	1615	14	US-10-331-907-3
3	8736	100.0	1615	15	US-10-464-368-82
4	8736	100.0	1615	16	US-10-789-378-50
5	8724	99.8	1627	13	US-10-087-192-1410
6	8715	99.7	1615	15	US-10-374-979-3
7	8715	99.7	1615	15	US-10-182-936A-3
8	8715	99.7	1615	16	US-10-731-739-3
9	8706	99.6	1615	15	US-10-374-979-4
10	8706	99.6	1615	15	US-10-182-936A-4
11	8706	99.6	1615	16	US-10-731-739-4
12	8705	99.6	1611	15	US-10-464-368-81
13	8607	98.5	1591	14	US-10-331-907-4

Sequence 43, Appl
Sequence 29, Appl
Sequence 39, Appl
Sequence 2, Appl
Sequence 42, Appl
Sequence 80, Appl
Sequence 75, Appl
Sequence 94, Appl
Sequence 44, Appl
Sequence 25, Appl
Sequence 84, Appl
Sequence 83, Appl
Sequence 1612, Ap
Sequence 1286, Ap
Sequence 1407, Ap
Sequence 7, Appl
Sequence 2, Appl
Sequence 711, App
Sequence 711, App
Sequence 6, Appl
Sequence 335, App
Sequence 335, App
Sequence 60, Appl
Sequence 5, Appl
Sequence 8, Appl
Sequence 6, Appl
Sequence 1723, Ap
Sequence 332, App
Sequence 333, App
Sequence 67, Appl
Sequence 71, Appl

ALIGNMENTS

RESULT 1

US-09-931-375A-2
; Sequence 2, Application US/09931375A
; Publication No. US20030027151A1
; GENERAL INFORMATION:
; APPLICANT: WARMAN, Matthew L.
; APPLICANT: GONG, Yaogin
; APPLICANT: OLSEN, Bjorn R.
; APPLICANT: ROMANI, Georges
; APPLICANT: ROMAN-ROMAN, Sergio
; TITLE OF INVENTION: REGULATOR GENE AND SYSTEM USEFUL FOR THE DIAGNOSIS AND THERAPY OF
; FILE REFERENCE: 38464-0004
; CURRENT APPLICATION NUMBER: US/09/931,375A
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/304,851
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/234,337
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/226,119
; PRIOR FILING DATE: 2000-08-18
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-931-375A-2

Query Match 100.0%; Score 8740; DB 10; Length 1615;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1615; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MEAAPGPPWPLLLLLLLLLLALCCGCPAPAAAPLFFANRRDRLVLDAGGVKLESTIVVS 60
Db 1 MEAAPGPPWPLLLLLLLLLLALCCGCPAPAAAPLFFANRRDRLVLDAGGVKLESTIVVS 60

QY 61 GLEDAADVDFQSKGAVYVTDVSEBIAIKOTYLNQTCGAQVQNVVVISGLVSPDGLACDWGK 120
DB 61 GLEDAADVDFQSKGAVYVTDVSEBIAIKOTYLNQTCGAQVQNVVVISGLVSPDGLACDWGK 120
QY 121 KLYWTDSETRIEVANLNGTSKVLFWQDLDOPRAIALDPAHGYMTDWTGTPRIERAG 180
DB 121 KLYWTDSETRIEVANLNGTSKVLFWQDLDOPRAIALDPAHGYMTDWTGTPRIERAG 180
QY 181 MDGSTRKIIIVSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPFRQKVVEGSLTHP 240
DB 181 MDGSTRKIIIVSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPFRQKVVEGSLTHP 240
QY 241 FALTLSGDTLYWTDQTRSIHACNKRKTKGKREIISALYSFMDIOVLQSOEROPFFHTTRCE 300
DB 241 FALTLSGDTLYWTDQTRSIHACNKRKTKGKREIISALYSFMDIOVLQSOEROPFFHTTRCE 300
QY 301 EDNGGCSHLCLLSPSEPFYTCAPTGVOLQDNGRCTCKAGAEVLLARLTDLRRISLDTTP 360
DB 301 EDNGGCSHLCLLSPSEPFYTCAPTGVOLQDNGRCTCKAGAEVLLARLTDLRRISLDTTP 360
QY 361 DFTDIVLQVDDIRHAIADIDPLEGVYVWTDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
DB 361 DFTDIVLQVDDIRHAIADIDPLEGVYVWTDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
QY 421 IAVDWARNLYWTDGTDRIEVTRLNGTSRKILVSEDLDEPRALHVPVGLMYWTDWGE 480
DB 421 IAVDWARNLYWTDGTDRIEVTRLNGTSRKILVSEDLDEPRALHVPVGLMYWTDWGE 480
QY 481 NPKIECANLDQERRVLVNASLWPNGLALDIOEGKLYWDAKTDKIEVINVDGTKRRTL 540
DB 481 NPKIECANLDQERRVLVNASLWPNGLALDIOEGKLYWDAKTDKIEVINVDGTKRRTL 540
QY 541 LEDKLPHPFGTLLGDFIYVTDWQRRSIEVHKVKAARDVIIDQLPDLMLGLKAVNAVAV 600
DB 541 LEDKLPHPFGTLLGDFIYVTDWQRRSIEVHKVKAARDVIIDQLPDLMLGLKAVNAVAV 600
QY 601 GTNPCADRNGGCSHLCTFTPHATRCGPIGLELLSDMTKICVPEAFVFTSRAAHLRLSL 660
DB 601 GTNPCADRNGGCSHLCTFTPHATRCGPIGLELLSDMTKICVPEAFVFTSRAAHLRLSL 660
QY 661 ETNNNDVAIPTGVKEASALDPDVSNHHIYVTDVSLKTIISRAFMNGSSVEHVFEGLDYP 720
DB 661 ETNNNDVAIPTGVKEASALDPDVSNHHIYVTDVSLKTIISRAFMNGSSVEHVFEGLDYP 720
QY 721 EGMADVWGNLYWADTCTGNRIEVARLDGQROVLVWDLNPRSLALDPTKGYIYWTW 780
DB 721 EGMADVWGNLYWADTCTGNRIEVARLDGQROVLVWDLNPRSLALDPTKGYIYWTW 780
QY 781 GGPRIVRAFMVDMGTNCMTLVKVRANDLTIDYADQRLYVTDLTNTMISSNMLGQERVV 840
DB 781 GGPRIVRAFMVDMGTNCMTLVKVRANDLTIDYADQRLYVTDLTNTMISSNMLGQERVV 840
QY 841 IADDLPHFPGLTQVSDIYVTDWNLHLSIERADKTSGRNRTLIQGHLPFVMDILVPHSSRQ 900
DB 841 IADDLPHFPGLTQVSDIYVTDWNLHLSIERADKTSGRNRTLIQGHLPFVMDILVPHSSRQ 900
QY 901 DGLNDCHNNGCQGLCIAIPGGHRCGASHYTLDPSSRNCSPPTTLLFSQKSAISRM 960
DB 901 DGLNDCHNNGCQGLCIAIPGGHRCGASHYTLDPSSRNCSPPTTLLFSQKSAISRM 960
QY 961 PDDQHSPLIILPLHGLRNVAKIDVDPDKFYVWVDGRQNIKRAKDDGTQPPVLTSLSQGG 1020
DB 961 PDDQHSPLIILPLHGLRNVAKIDVDPDKFYVWVDGRQNIKRAKDDGTQPPVLTSLSQGG 1020
QY 1021 NPDQPHDLSIDISRTLFTWCEATNTINVRHLSGEAMGVVLRGRDPRKRAIVVNAERG 1080
DB 1021 NPDQPHDLSIDISRTLFTWCEATNTINVRHLSGEAMGVVLRGRDPRKRAIVVNAERG 1080
QY 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVVALVVDNTLGLFWVDADAKRIESCD 1140
DB 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVVALVVDNTLGLFWVDADAKRIESCD 1140

QY 1141 LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG 1200
DB 1141 LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG 1200
QY 1201 IHAVEEVSLEBSAHPCARDNGGCSHICIAKGDGTPRCSCPVLHVLVLLQNLITCGSPPTCS 1260
DB 1201 IHAVEEVSLEBSAHPCARDNGGCSHICIAKGDGTPRCSCPVLHVLVLLQNLITCGSPPTCS 1260
QY 1261 PDQFACATGEIDCIPGAWRCDFPECCDDQSDDEEGCPVCSAAQFPCCARQCQVDLRLRCDE 1320
DB 1261 PDQFACATGEIDCIPGAWRCDFPECCDDQSDDEEGCPVCSAAQFPCCARQCQVDLRLRCDE 1320
QY 1321 ADCQDRSEADCAICLNPQRCASGQVLKQOCDSPDCIDGSDDELMLCMTITPPSDDS 1380
DB 1321 ADCQDRSEADCAICLNPQRCASGQVLKQOCDSPDCIDGSDDELMLCMTITPPSDDS 1380
QY 1381 PAHSSAIGPVGIIILSLFVWGGVYVFCORVVCQRYAGANGPPHYYVSGTDPHVLNFTAP 1440
DB 1381 PAHSSAIGPVGIIILSLFVWGGVYVFCORVVCQRYAGANGPPHYYVSGTDPHVLNFTAP 1440
QY 1441 GGSQHGPTGTIACGKSMSSVSLMGRGGVFLYDRNHVTGASSSSSSSTKATLYPPIINP 1500
DB 1441 GGSQHGPTGTIACGKSMSSVSLMGRGGVFLYDRNHVTGASSSSSSSTKATLYPPIINP 1500
QY 1501 PPSPATDPSLYNMDMFYSSNIPATARPYPYIIIRGMAPPPTTPTCSTDVCDSDYSASRWKAS 1560
DB 1501 PPSPATDPSLYNMDMFYSSNIPATARPYPYIIIRGMAPPPTTPTCSTDVCDSDYSASRWKAS 1560
QY 1561 KYIYLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
DB 1561 KYIYLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 2

US-10-331-907-3

; Sequence 3, Application us/10331907

; Publication No. US20030181660A1

; GENERAL INFORMATION:

; APPLICANT: Todd, John A

; Hess, John W

; Caskey, Charles T

; Cox, Roger D

; Gerhold, David

; Hammond, Holly

; Hey, Patricia

; Kawaguchi, Yoshihiko

; Merriman, Tony R

; Metzker, Michael L

; TITLE OF INVENTION: No. US20030181660A1e1 LDL-Receptor

; NUMBER OF SEQUENCES: 455

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Nixon and Vanderhye

; STREET: 1100 No. US20030181660A1e1 Glebe Road, Eighth Floor

; CITY: Arlington

; STATE: Virginia

; COUNTRY: US

; ZIP: VA 22201-4714

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/331,907

; FILING DATE: 31-Dec-2002

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/402,923A

; FILING DATE: 14-Feb-2001

; APPLICATION NUMBER: PCT/GB98/01102

; FILING DATE: 15-APR-1998

; APPLICATION NUMBER: US 60/043,553

; FILING DATE: 15-APR-1997

; APPLICATION NUMBER: US 60/048,740

```
;
; FILING DATE: 05-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: B. J. Sadoff
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 620-81
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4091
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1615 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-331-907-3

Query Match      100.0%; Score 8736; DB 14; Length 1615;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEAAPGPPPLLLLLLLLLALCCPAPAAAPLFFANRRDRLVLDAGGVKLESTIVS 60
DB 1 MEAAPGPPPLLLLLLLLLALCCPAPAAAPLFFANRRDRLVLDAGGVKLESTIVS 60
QY 61 GLEDAADVDFQFSKGVYTDVSEEAIKQTYLNQTAGAQQNVVLSGLVSPDGLACDVGK 120
DB 61 GLEDAADVDFQFSKGVYTDVSEEAIKQTYLNQTAGAQQNVVLSGLVSPDGLACDVGK 120
QY 121 KLYWTDSETRIEVANLNGTSRKVLFWQDLDPRAIALDPAGHYMYTWDGNETPRIERAG 180
DB 121 KLYWTDSETRIEVANLNGTSRKVLFWQDLDPRAIALDPAGHYMYTWDGNETPRIERAG 180
QY 181 MDGSTRKIIIVDSDIYWPNGLTIDLEOKLYWADAKLSFIHRANLDGFRQKVVSGSLTTP 240
DB 181 MDGSTRKIIIVDSDIYWPNGLTIDLEOKLYWADAKLSFIHRANLDGFRQKVVSGSLTTP 240
QY 241 FALTLSDGTLTYTDWQTRSIIHACNKRKGKRKEILSALYSPMDIQVLSQERQPFHTTRCE 300
DB 241 FALTLSDGTLTYTDWQTRSIIHACNKRKGKRKEILSALYSPMDIQVLSQERQPFHTTRCE 300
QY 301 EDNGGCSHLCLLSPSEPFYTCACPTGVLQNGRTCKAGAEVLLARRDLRRLISLDTTP 360
DB 301 EDNGGCSHLCLLSPSEPFYTCACPTGVLQNGRTCKAGAEVLLARRDLRRLISLDTTP 360
QY 361 DFTDIVLQVDDIRHAIADYDPLEGYVYTDDEVRARRAYLDGSGAQTILVNTINDPDG 420
DB 361 DFTDIVLQVDDIRHAIADYDPLEGYVYTDDEVRARRAYLDGSGAQTILVNTINDPDG 420
QY 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALHPVLMGLMYTWDGWE 480
DB 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALHPVLMGLMYTWDGWE 480
QY 481 NPKIECANLQGERVRLVNASLGNWPNGLALDLQEGKLYWDAKTDKIEVINVDGTKRRTL 540
DB 481 NPKIECANLQGERVRLVNASLGNWPNGLALDLQEGKLYWDAKTDKIEVINVDGTKRRTL 540
QY 541 LEDKLPHIFGTLTGDFIYTDWQRRSIEVRHVKKASRDVLIIDQLPDLMLGLKANVAKV 600
DB 541 LEDKLPHIFGTLTGDFIYTDWQRRSIEVRHVKKASRDVLIIDQLPDLMLGLKANVAKV 600
QY 601 GTNPNADRNGGCSHLCEFTPHATRCGCPIGLELLSDMKTCIVPEAFVLTFSRAAHRISL 660
DB 601 GTNPNADRNGGCSHLCEFTPHATRCGCPIGLELLSDMKTCIVPEAFVLTFSRAAHRISL 660
QY 661 ETNNNDVAIPLTGKVASALDFVSNHHIYWTVDLSLTKISRAFPNGSSVEHVVEFGLDYP 720
DB 661 ETNNNDVAIPLTGKVASALDFVSNHHIYWTVDLSLTKISRAFPNGSSVEHVVEFGLDYP 720
QY 721 EGMADVAMGNLYWADGTNRRIEVARLDGQPROVLVWRDLNPRSLALDPTKGIYIYTW 780
DB 721 EGMADVAMGNLYWADGTNRRIEVARLDGQPROVLVWRDLNPRSLALDPTKGIYIYTW 780
QY 781 GSKPRIIVRAFMDGTNCMTLVKVRANDLTIDYADQRLYTDLTDTNMISSNMLQSERVV 840
DB 781 GSKPRIIVRAFMDGTNCMTLVKVRANDLTIDYADQRLYTDLTDTNMISSNMLQSERVV 840
QY 841 IADDLPHPPFGLTQYSDIYIYTDWNLHLSIERADKTSGRNRTLIQGHLDFTWMDILVPHSSRQ 900
DB 841 IADDLPHPPFGLTQYSDIYIYTDWNLHLSIERADKTSGRNRTLIQGHLDFTWMDILVPHSSRQ 900
QY 901 DGLNDCMHNNGCCQQLCLAIIPGHRCCGASHYTLDPSSRNCSPPTTFLFQSKSAISMI 960
DB 901 DGLNDCMHNNGCCQQLCLAIIPGHRCCGASHYTLDPSSRNCSPPTTFLFQSKSAISMI 960
QY 961 PDDQHSPLDILPLHLGLRNVAIDYDPLDKFIYWDGQRNIKRAKDDGTQPFVLTLSLQSQ 1020
DB 961 PDDQHSPLDILPLHLGLRNVAIDYDPLDKFIYWDGQRNIKRAKDDGTQPFVLTLSLQSQ 1020
QY 1021 NPDQPHDLSIDIYSRTILFWTCEATNTINVHRLSGEAMGVVLRGDRDPRAIIVNAERG 1080
DB 1021 NPDQPHDLSIDIYSRTILFWTCEATNTINVHRLSGEAMGVVLRGDRDPRAIIVNAERG 1080
QY 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESC 1140
DB 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESC 1140
QY 1141 LSGANRLTLEDANIVQPLGLTILGKLYWIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG 1200
DB 1141 LSGANRLTLEDANIVQPLGLTILGKLYWIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG 1200
QY 1201 IHAVEEVSLEBSFSAHPCARDNGGCSHICIAKGDGTGPRCSPVHLVLLQNLTCGEPPTCS 1260
DB 1201 IHAVEEVSLEBSFSAHPCARDNGGCSHICIAKGDGTGPRCSPVHLVLLQNLTCGEPPTCS 1260
QY 1261 PDQFACATGETDCIPGAWRCDFPECDQSDDEECPCVCSAAQFPCARQCVDLRLRCDE 1320
DB 1261 PDQFACATGETDCIPGAWRCDFPECDQSDDEECPCVCSAAQFPCARQCVDLRLRCDE 1320
QY 1321 ADCQDRSDEADCDACILPNQFRCASGQCVLIKQCCDSFPDCIDGSDDELWCBITPPSDDS 1380
DB 1321 ADCQDRSDEADCDACILPNQFRCASGQCVLIKQCCDSFPDCIDGSDDELWCBITPPSDDS 1380
QY 1381 PAHSSAIGPVIGIILSLFVNGGVYFVQVVCQRYAGANGPPHYVSGTHVPLNFIAP 1440
DB 1381 PAHSSAIGPVIGIILSLFVNGGVYFVQVVCQRYAGANGPPHYVSGTHVPLNFIAP 1440
QY 1441 GGSQHGPFPTGIACQSKMSSVSLMGGRGVPLYDRNHVTGASSSSSSSTKATLPPILNP 1500
DB 1441 GGSQHGPFPTGIACQSKMSSVSLMGGRGVPLYDRNHVTGASSSSSSSTKATLPPILNP 1500
QY 1501 PPSPATDPSLYNMDMFYSSNIPATAPRYPYIIRGMAPPPTTTCSTDVCDSDYSASRWKAS 1560
DB 1501 PPSPATDPSLYNMDMFYSSNIPATAPRYPYIIRGMAPPPTTTCSTDVCDSDYSASRWKAS 1560
QY 1561 KYILDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
DB 1561 KYILDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 3
US-10-464-368-82
; Sequence 82, Application US/10464368
; Publication No. US20040023356A1
; GENERAL INFORMATION:
; APPLICANT: Krumlauf, Robb
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
; FILE REFERENCE: 40716-IP-017
; CURRENT APPLICATION NUMBER: US/10/464,368
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/388,970
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 82
; LENGTH: 1615
```

```

; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-10-464-368-82

Query Match      100.0%; Score 8736; DB 15; Length 1615;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEAAAPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVS 60
DB 1 MEAAAPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVS 60
QY 61 GLEDAAAVDQFQSKGAVYWDVSEBAIKOTYLNQTAAGVQNVVLSGLVSPDGLACDVGK 120
DB 61 GLEDAAAVDQFQSKGAVYWDVSEBAIKOTYLNQTAAGVQNVVLSGLVSPDGLACDVGK 120
QY 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLPRAIALDPAGHYMYWTDWGETPRIERAG 180
DB 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLPRAIALDPAGHYMYWTDWGETPRIERAG 180
QY 181 MDGSTRKIIIVSDIYWPNGLTIDLEOKLYWADAKLSFTHRANLDGSPRQKVVEGSLTHP 240
DB 181 MDGSTRKIIIVSDIYWPNGLTIDLEOKLYWADAKLSFTHRANLDGSPRQKVVEGSLTHP 240
QY 241 FALTLSGDTLYWTDWQTRSIHACNKRKTGKRKEILSALYSMDIOVLQSOERQPPHTRCE 300
DB 241 FALTLSGDTLYWTDWQTRSIHACNKRKTGKRKEILSALYSMDIOVLQSOERQPPHTRCE 300
QY 301 EDNGGCSHCLLSPSEPFYTCAPTGVQIQDNNGRTCKAGABEVLLARRTDLRLISLDTLP 360
DB 301 EDNGGCSHCLLSPSEPFYTCAPTGVQIQDNNGRTCKAGABEVLLARRTDLRLISLDTLP 360
QY 361 DFTDVLQVDDIRHAIADYDPLEGYVYWDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
DB 361 DFTDVLQVDDIRHAIADYDPLEGYVYWDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
QY 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRATALHPVMGLMYWTDWGE 480
DB 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRATALHPVMGLMYWTDWGE 480
QY 481 NPKIECANLDGQERRVLVNASIGWPNGLALDQEGKLYWDAKTDKIEVINVDGTKRTL 540
DB 481 NPKIECANLDGQERRVLVNASIGWPNGLALDQEGKLYWDAKTDKIEVINVDGTKRTL 540
QY 541 LEDKLPHIFGFTLLGDFIYWDWQRRSIRERHVKYKASRDVLIIDQLPDLMLGLKAVNAVW 600
DB 541 LEDKLPHIFGFTLLGDFIYWDWQRRSIRERHVKYKASRDVLIIDQLPDLMLGLKAVNAVW 600
QY 601 GTNPCADRNGGCSHLCFTTPHATRCGCPGLELLSDMKTCIVPEAFVFTSRAAIHRISL 660
DB 601 GTNPCADRNGGCSHLCFTTPHATRCGCPGLELLSDMKTCIVPEAFVFTSRAAIHRISL 660
QY 661 ETNNNDVAIPUTGVKEASALDFVSNHHIYWDVSLKTSIRAFWNGSSVEHVVEFGLDYP 720
DB 661 ETNNNDVAIPUTGVKEASALDFVSNHHIYWDVSLKTSIRAFWNGSSVEHVVEFGLDYP 720
QY 721 EGMADVWNGKLYWADTGTNRIEVARLDGQPROVLVWRDLNPRSLALDPTKGIYIYTEW 780
DB 721 EGMADVWNGKLYWADTGTNRIEVARLDGQPROVLVWRDLNPRSLALDPTKGIYIYTEW 780
QY 781 GCKPRIVRAAFMDGNTCMVLVDKVRANDLTIDYADQRLYTDLDTNMISSNMLGQERVV 840
DB 781 GCKPRIVRAAFMDGNTCMVLVDKVRANDLTIDYADQRLYTDLDTNMISSNMLGQERVV 840
QY 841 IADLPHFPFGITQYSDIYIYTDWNLHSTIERADKTSGRNRTLIQGHLDPMVILVPHSSRQ 900
DB 841 IADLPHFPFGITQYSDIYIYTDWNLHSTIERADKTSGRNRTLIQGHLDPMVILVPHSSRQ 900
QY 901 DGLNDCHNNGCCQOLCLAIPEGHRCGCASHYTLDPSSNCSPTTFLIFSOKAISRM 960
DB 901 DGLNDCHNNGCCQOLCLAIPEGHRCGCASHYTLDPSSNCSPTTFLIFSOKAISRM 960
QY 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFYWVDGRQNIKRAKDDGTPPVLTSLSQGG 1020
DB 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFYWVDGRQNIKRAKDDGTPPVLTSLSQGG 1020
QY 1021 NPDRPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLGRDRKPRAIIVNAERG 1080
DB 1021 NPDRPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLGRDRKPRAIIVNAERG 1080
QY 1081 LYFTNMQRRAAKIERAALDGTREVLFTTGLIRPVVALVVDNTLGLFWVDADLKRIESCD 1140
DB 1081 LYFTNMQRRAAKIERAALDGTREVLFTTGLIRPVVALVVDNTLGLFWVDADLKRIESCD 1140
QY 1141 LSGANRLTLEDAIVQPLGLTILGKHLWIDRQOQMIERVEKTTGDKETRIQGRVAHLTG 1200
DB 1141 LSGANRLTLEDAIVQPLGLTILGKHLWIDRQOQMIERVEKTTGDKETRIQGRVAHLTG 1200
QY 1201 IHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSCPVLHVLVLLQNLITTCGEPPTCS 1260
DB 1201 IHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSCPVLHVLVLLQNLITTCGEPPTCS 1260
QY 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEEGCPVCSAAQFPFCARGQCVDLRLRCDGE 1320
DB 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEEGCPVCSAAQFPFCARGQCVDLRLRCDGE 1320
QY 1321 ADCQDRSDEADCDALCLPNQPRCASGOCVLIKQOCDSFPDCIDGSDDELWCEITKPPSDDS 1380
DB 1321 ADCQDRSDEADCDALCLPNQPRCASGOCVLIKQOCDSFPDCIDGSDDELWCEITKPPSDDS 1380
QY 1381 PAHSSAIGPVIIGIILSLFVMGVYFVQVRVVCQRYAGANGPFPHEVYSGTTHVPLNFIAP 1440
DB 1381 PAHSSAIGPVIIGIILSLFVMGVYFVQVRVVCQRYAGANGPFPHEVYSGTTHVPLNFIAP 1440
QY 1441 GGSQHGFTGTIACGKSMSSVSLMGRGGVPLYDRNHVGTGASSSSSSSTKATLYPPILNP 1500
DB 1441 GGSQHGFTGTIACGKSMSSVSLMGRGGVPLYDRNHVGTGASSSSSSSTKATLYPPILNP 1500
QY 1501 PPSPATPSLYNMDMEYSSNIPATRPVPIIRGMAPPPTPCSTDYCDSDYSASRWKAS 1560
DB 1501 PPSPATPSLYNMDMEYSSNIPATRPVPIIRGMAPPPTPCSTDYCDSDYSASRWKAS 1560
QY 1561 KYIYLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLFPFPPSPCTDSS 1615
DB 1561 KYIYLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLFPFPPSPCTDSS 1615

RESULT 4
US-10-789-378-50
; Sequence 50, Application US/10789378
; Publication No. US20050003390A1
; GENERAL INFORMATION:
; APPLICANT: Azenovich, Sergey
; APPLICANT: Stull, Robert
; APPLICANT: Gelman, Marina
; APPLICANT: Chui, Kitty
; APPLICANT: Ng, Dean
; TITLE OF INVENTION: DIAGNOSTIC METHODS FOR CANCER DETECTION
; FILE REFERENCE: 5189-2
; CURRENT APPLICATION NUMBER: US/10789,378
; CURRENT FILING DATE: 2004-02-26
; PRIOR APPLICATION NUMBER: 10/441925
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: 60/381619
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/450886
; PRIOR FILING DATE: 2003-02-26
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-789-378-50

Query Match      100.0%; Score 8736; DB 16; Length 1615;
```

Best Local Similarity 99.9%; Pred. No. 0; Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
Qy	1	MEAAPGPPWPLLLLLLLLLLALCCCPAPAAASPLLLFANRREDVRLVDAGGVKLESTIVS	60						
Db	1	MEAAPGPPWPLLLLLLLLLLALCCCPAPAAASPLLLFANRREDVRLVDAGGVKLESTIVS	60						
Qy	61	GLEDAAAVDFSKGAVVTVDVSEAIKQTYLNQTGAOVNVIISGLVSPDGLACDWGK	120						
Db	61	GLEDAAAVDFSKGAVVTVDVSEAIKQTYLNQTGAOVNVIISGLVSPDGLACDWGK	120						
Qy	121	KLYTDSQTRNRIEVANLNGTSRKVLFWQDLPQRAIALDPAGHYMYTWDGQETPRIERAG	180						
Db	121	KLYTDSQTRNRIEVANLNGTSRKVLFWQDLPQRAIALDPAGHYMYTWDGQETPRIERAG	180						
Qy	181	MDGSTRKLIIVDSDIYWPNGLTIDLEBOKLYWADAKLSPIHEANLDGSPROKVGSLTTP	240						
Db	181	MDGSTRKLIIVDSDIYWPNGLTIDLEBOKLYWADAKLSPIHEANLDGSPROKVGSLTTP	240						
Qy	241	FALTLSGDTLYTWDQTRSIHACNKRITGGKKEILSALYSPMDIQVLSQERQPFHTRCE	300						
Db	241	FALTLSGDTLYTWDQTRSIHACNKRITGGKKEILSALYSPMDIQVLSQERQPFHTRCE	300						
Qy	301	EDNGGCSHLCLLSPSEPPYTCACPTGVLQNGRTCKAGAEVLLARRTDLRRISLDT	360						
Db	301	EDNGGCSHLCLLSPSEPPYTCACPTGVLQNGRTCKAGAEVLLARRTDLRRISLDT	360						
Qy	361	DFTDILVQVDDIRHAIADIDPLRGYVYTTDDEVRATRRAYLDGSGAQTLVNTINPDG	420						
Db	361	DFTDILVQVDDIRHAIADIDPLRGYVYTTDDEVRATRRAYLDGSGAQTLVNTINPDG	420						
Qy	421	IADVARNLYTDTGTDRIEVTRNGTSRKILVSIEDLDEPRALHPVGLMYTWDGGE	480						
Db	421	IADVARNLYTDTGTDRIEVTRNGTSRKILVSIEDLDEPRALHPVGLMYTWDGGE	480						
Qy	481	NPKIECANLDQERRVLVNASLGNPLALDQEGKLYWDAKTKIEVINVDGTKRTL	540						
Db	481	NPKIECANLDQERRVLVNASLGNPLALDQEGKLYWDAKTKIEVINVDGTKRTL	540						
Qy	541	LEDKLPHIFGTLTGDFTYTDWQRRSIEVHKYKASRDVLIIDQLPDLMLKANVAKV	600						
Db	541	LEDKLPHIFGTLTGDFTYTDWQRRSIEVHKYKASRDVLIIDQLPDLMLKANVAKV	600						
Qy	601	GTNPACDRNGCSHLCEFTTHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHISL	660						
Db	601	GTNPACDRNGCSHLCEFTTHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHISL	660						
Qy	661	ETNNVDVAIPLTGKESALDQFVSNHHIYWTDSVSLKTIISRAFNWGSVEHVVEFGLDYP	720						
Db	661	ETNNVDVAIPLTGKESALDQFVSNHHIYWTDSVSLKTIISRAFNWGSVEHVVEFGLDYP	720						
Qy	721	EGMAVDWGNLYWADTGTNRIEVARLDGQPRQLVWRDLNPSRLALDPTKGYIYTW	780						
Db	721	EGMAVDWGNLYWADTGTNRIEVARLDGQPRQLVWRDLNPSRLALDPTKGYIYTW	780						
Qy	781	GGKPRIIVRAFMDGTNCTLVKVRANDLTIDYADQRLYMTDLTNIESNNMLGQSRVV	840						
Db	781	GGKPRIIVRAFMDGTNCTLVKVRANDLTIDYADQRLYMTDLTNIESNNMLGQSRVV	840						
Qy	841	IADLPHFPGLTQVSDYIYWTDMNLHSIERADKTSGRNRTLIOGHLDPMVILVFHSSRQ	900						
Db	841	IADLPHFPGLTQVSDYIYWTDMNLHSIERADKTSGRNRTLIOGHLDPMVILVFHSSRQ	900						
Qy	901	DGLNDCMHNNGCCQOLCLAPGGHRCGCASHYTLDPSSRNCSPPTTLLFSQKSAISRM	960						
Db	901	DGLNDCMHNNGCCQOLCLAPGGHRCGCASHYTLDPSSRNCSPPTTLLFSQKSAISRM	960						
Qy	961	PDDQHSPLILPLHGLNRKVIDYDPLDKFYWDGRQNIKRAKDDGTQPPVLTSLSQG	1020						
Db	961	PDDQHSPLILPLHGLNRKVIDYDPLDKFYWDGRQNIKRAKDDGTQPPVLTSLSQG	1020						
Qy	1021	NPDRQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERGY	1080						

Db	1021	NPDRQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERGY	1080						
Qy	1081	LYFTNMQDRAKIERAALDGTREVLFTTGLIRVALVVDNTLGLFWVDADLKRIESCD	1140						
Db	1081	LYFTNMQDRAKIERAALDGTREVLFTTGLIRVALVVDNTLGLFWVDADLKRIESCD	1140						
Qy	1141	LSGANRLTLEDANIVQPLGLTILGKHYIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG	1200						
Db	1141	LSGANRLTLEDANIVQPLGLTILGKHYIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG	1200						
Qy	1201	IHAVEEVSLEBFSAPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQNLLTCGEPPTCS	1260						
Db	1201	IHAVEEVSLEBFSAPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQNLLTCGEPPTCS	1260						
Qy	1261	PQFACATGETDCIPGAWRCDFPECDQSDDEECPCVCSAAQFPCCARGCQVDLRLRCDGE	1320						
Db	1261	PQFACATGETDCIPGAWRCDFPECDQSDDEECPCVCSAAQFPCCARGCQVDLRLRCDGE	1320						
Qy	1321	ADCDRDEADCDACILPNQFRCASGQCVLTKQCCDSFPDCIDGSDDELMCEBITKPPSDDS	1380						
Db	1321	ADCDRDEADCDACILPNQFRCASGQCVLTKQCCDSFPDCIDGSDDELMCEBITKPPSDDS	1380						
Qy	1381	PAHSSAIGPVIGIILSLFVWGCVYFVQVQVQRYAGANGPFPHEYVSGTTPHVLNFTAP	1440						
Db	1381	PAHSSAIGPVIGIILSLFVWGCVYFVQVQVQRYAGANGPFPHEYVSGTTPHVLNFTAP	1440						
Qy	1441	GGSHGPGFTGIACGKSNMSSVSLMGGRGVPLYDRNHVGTGASSSSSTKATLYPPLNP	1500						
Db	1441	GGSHGPGFTGIACGKSNMSSVSLMGGRGVPLYDRNHVGTGASSSSSTKATLYPPLNP	1500						
Qy	1501	PPSPATDPSLYNMDMFYSSNIPATPYRPIYIRGMAPPTTPTCSTDVDCSDYSASRWKAS	1560						
Db	1501	PPSPATDPSLYNMDMFYSSNIPATPYRPIYIRGMAPPTTPTCSTDVDCSDYSASRWKAS	1560						
Qy	1561	KYILDNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS	1615						
Db	1561	KYILDNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS	1615						
RESULT 5									
US-10-087-192-1410									
; Sequence 1410, Application US/10087192									
; Publication No. US20020182586A1									
; GENERAL INFORMATION:									
; APPLICANT: Morris, David W.									
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR									
; FILE REFERENCE: 529452000122									
; CURRENT APPLICATION NUMBER: US/10/087,192									
; CURRENT FILING DATE: 2002-03-01									
; PRIOR APPLICATION NUMBER: US 09/747,377									
; PRIOR FILING DATE: 2000-12-22									
; PRIOR APPLICATION NUMBER: US 09/798,586									
; PRIOR FILING DATE: 2001-03-02									
; NUMBER OF SEQ ID NOS: 2059									
; SOFTWARE: FastSeq for Windows Version 4.0									
; SEQ ID NO 1410									
; LENGTH: 1627									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-10-087-192-1410									
Query Match 99.8%; Score 8724; DB 13; Length 1627;									
Best Local Similarity 99.3%; Pred. No. 0;									
Matches 1615; Conservative 0; Mismatches 0; Indels 12; Gaps 1;									
Qy	1	MEAAPGPPWPLLLLLLLLLLALCCCPAPAAASPLLLFANRREDVRLVDAGGVKLESTIVS	60						
Db	1	MEAAPGPPWPLLLLLLLLLLALCCCPAPAAASPLLLFANRREDVRLVDAGGVKLESTIVS	60						
Qy	61	GLEDAAAVDFSKGAVVTVDVSEAIKQTYLNQTGAOVNVIISGLVSPDGLACDWGK	120						

Db 61 GLEDAAAVDQFSGKGVYWTDSVEEAIKQTYLNTQTGAAGVQNVVLSGLVSPDGLACDWGK 120
Qy 121 KLYWTDSETRLEVANLNGTSKRVFWQDLDPRAIALDPAGYMYWTWDGTPRIERAG 180
Db 121 KLYWTDSETRLEVANLNGTSKRVFWQDLDPRAIALDPAGYMYWTWDGTPRIERAG 180
Qy 181 MDGSTRKIIIVDSDIYWPNGLTLDLEOKLYWADAKLSPIHRLANLNGSFRQKVEGSLTHP 240
Db 181 MDGSTRKIIIVDSDIYWPNGLTLDLEOKLYWADAKLSPIHRLANLNGSFRQKVEGSLTHP 240
Qy 241 FALTLSGDTLYWTDQTRSIIHACNKRRTGKRKEILSALYSFMDIOVLQOEROPFFHTRCE 300
Db 241 FALTLSGDTLYWTDQTRSIIHACNKRRTGKRKEILSALYSFMDIOVLQOEROPFFHTRCE 300
Qy 301 EDNGGCSHLCLSPSEPFYTCAPTGVQLQDNGRTCKAGAEVLLIARRDLRRLSLDTP 360
Db 301 EDNGGCSHLCLSPSEPFYTCAPTGVQLQDNGRTCKAGAEVLLIARRDLRRLSLDTP 360
Qy 361 DFTDIVLQVDDIRHAIADIDYDPLEGVVYWTDEVRARAYLDGSGAOTLVNTEINDPDG 420
Db 361 DFTDIVLQVDDIRHAIADIDYDPLEGVVYWTDEVRARAYLDGSGAOTLVNTEINDPDG 420
Qy 421 IAVDVARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALIALHPVMGLMYWTDWGE 480
Db 421 IAVDVARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALIALHPVMGLMYWTDWGE 480
Qy 481 NPKIECANLDGERRVLVNASLGNGLALDLOEGLYWDGAKTKIEVINVDGTRKRTL 540
Db 481 NPKIECANLDGERRVLVNASLGNGLALDLOEGLYWDGAKTKIEVINVDGTRKRTL 540
Qy 541 LEDKLPHIFGFTLLGDFIYWTWQRRSIERVHKVKSARDVIIQDLPDLMLKAVNAVVKV 600
Db 541 LEDKLPHIFGFTLLGDFIYWTWQRRSIERVHKVKSARDVIIQDLPDLMLKAVNAVVKV 600
Qy 601 GTPNCDNRGGCSHLCTFFTHATRCGCPGLELLSDMKTCIVPEAFVFTSRAAIHRLSL 660
Db 601 GTPNCDNRGGCSHLCTFFTHATRCGCPGLELLSDMKTCIVPEAFVFTSRAAIHRLSL 660
Qy 661 ETNNNDVAIPLTGVKEASALDFDVSNHHIYWTDSLTKTISRAFPMNGSSVEHVVERGLDYP 720
Db 661 ETNNNDVAIPLTGVKEASALDFDVSNHHIYWTDSLTKTISRAFPMNGSSVEHVVERGLDYP 720
Qy 721 EGMAVDWNGKLYWADTGTNRLEVARLDGQFQVLRWDLNPRSLALDPTKGYLYWTEW 780
Db 721 EGMAVDWNGKLYWADTGTNRLEVARLDGQFQVLRWDLNPRSLALDPTKGYLYWTEW 780
Qy 781 GSKPRIVAFMDGTNCMTLVKVGSRANDITIDYADQRLYWTDLDTNMTIESSNMLQOERVV 840
Db 781 GSKPRIVAFMDGTNCMTLVKVGSRANDITIDYADQRLYWTDLDTNMTIESSNMLQOERVV 840
Qy 841 IADDLPHFPGLTOYSDIYWTWDLNLSIERADKTSGRNRTLIQGLHDFVMDILVPHSSRQ 900
Db 841 IADDLPHFPGLTOYSDIYWTWDLNLSIERADKTSGRNRTLIQGLHDFVMDILVPHSSRQ 900
Qy 901 DGLNDCMHNGCGQLCLAIPEGHRCGCASHYTLDPSSRNCSPPTFFLLFSQKSAISRMI 960
Db 901 DGLNDCMHNGCGQLCLAIPEGHRCGCASHYTLDPSSRNCSPPTFFLLFSQKSAISRMI 960
Qy 961 PDDQHSPLILPLHGLRNVKAIDYDPLDKFIYWDGGRONIKRAKDDGT----- 1008
Db 961 PDDQHSPLILPLHGLRNVKAIDYDPLDKFIYWDGGRONIKRAKDDGT----- 1008
Qy 1009 QPFLVTLSSQGNPDROPHDLSIDYSLTFTWTCENTNTINVRHLSGEAMGVLRGRDCK 1068
Db 1021 QPFLVTLSSQGNPDROPHDLSIDYSLTFTWTCENTNTINVRHLSGEAMGVLRGRDCK 1080
Qy 1069 PRAIVVNAERGVLFTNNQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKLF 1128
Db 1081 PRAIVVNAERGVLFTNNQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKLF 1140
Qy 1129 VDADLKRIESCDSLGNRLTLEDANI VQPLGTLILGKHLWIDRQOQMIERVEKTTGDGR 1188
Db 1141 VDADLKRIESCDSLGNRLTLEDANI VQPLGTLILGKHLWIDRQOQMIERVEKTTGDGR 1200

Qy 1189 TRIQGRVAHLGTGIIHAVEVSLSEEPSAHPCARDNGGCSHCIAKGDGTPRCPCVHLVLLQ 1248
Db 1201 TRIQGRVAHLGTGIIHAVEVSLSEEPSAHPCARDNGGCSHCIAKGDGTPRCPCVHLVLLQ 1260
Qy 1249 NLLTCGEPTCSPOQFACATGEIDICIPGAWCDGPECCDDQSDERGCPCVCSAAQPPCARG 1308
Db 1261 NLLTCGEPTCSPOQFACATGEIDICIPGAWCDGPECCDDQSDERGCPCVCSAAQPPCARG 1320
Qy 1309 QCVDLRLRCDEADQCQDRSDEADDAICLPNQFRCASQCQVLIIKQCCDFPDCIDGSDDEL 1368
Db 1321 QCVDLRLRCDEADQCQDRSDEADDAICLPNQFRCASQCQVLIIKQCCDFPDCIDGSDDEL 1380
Qy 1369 MCEITKPSDDSPAHSSAIGVIGIILSLFVWGVIYVQCRVQCRVAGANGPPPEHYVS 1428
Db 1381 MCEITKPSDDSPAHSSAIGVIGIILSLFVWGVIYVQCRVQCRVAGANGPPPEHYVS 1440
Qy 1429 GTPHVPLNFIAPGSGQHPFTGIACGSKMMSVSLMGGRGVPLYDRNHVTGASSSSSSS 1488
Db 1441 GTPHVPLNFIAPGSGQHPFTGIACGSKMMSVSLMGGRGVPLYDRNHVTGASSSSSSS 1500
Qy 1489 TKATLYPILNPPSPATDPSLYNMDMFYSSNIPATAPRYPIYIRGMAPPTTTCSTDVC 1548
Db 1501 TKATLYPILNPPSPATDPSLYNMDMFYSSNIPATAPRYPIYIRGMAPPTTTCSTDVC 1560
Qy 1549 DSDYSASRWKASKYKYLINSDSDYPPTPHSQYLSAEDSCPPSPATERSYFHLFPPPP 1608
Db 1561 DSDYSASRWKASKYKYLINSDSDYPPTPHSQYLSAEDSCPPSPATERSYFHLFPPPP 1620
Qy 1609 SPCTDSS 1615
Db 1621 SPCTDSS 1627

RESULT 6

US-10-374-979-3
; Sequence 3, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 3
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-374-979-3

Query Match 99.7%; Score 8715; DB 15; Length 1615;

Best Local Similarity 99.7%; Pred. No. 0;
Matches 1610; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MEAAPPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVDAGGVKLESTIVVS 60

Db 1 MEAAPPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVDAGGVKLESTIVVS 60

Qy 61 GLEDAAAVDQFSGKGVYWTDSVEEAIKQTYLNTQTGAAGVQNVVLSGLVSPDGLACDWGK 120

Db 61 GLEDAAAVDQFSGKGVYWTDSVEEAIKQTYLNTQTGAAGVQNVVLSGLVSPDGLACDWGK 120


```
Qy 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDPQRAIALDPAHGYMYTWDGTPRIERAG 180
Db 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDPQRAIALDPAHGYMYTWDGTPRIERAG 180
Qy 181 MDGSTRKLIIVSDIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPRQKVVGSLTHP 240
Db 181 MDGSTRKLIIVSDIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPRQKVVGSLTHP 240
Qy 241 FALTSLGDTLWTDQWTRSHACNKRGTGKKEILSALYSMDIQVLISOERQPFPHTRCE 300
Db 241 FALTSLGDTLWTDQWTRSHACNKRGTGKKEILSALYSMDIQVLISOERQPFPHTRCE 300
Qy 301 EDNGGCSHLCLLSPSEPPYTCACPTGVOLQNGRTCKAGAEVLLARTRDLRLSLDTP 360
Db 301 EDNGGCSHLCLLSPSEPPYTCACPTGVOLQNGRTCKAGAEVLLARTRDLRLSLDTP 360
Qy 361 DPTDILVQVDDIRHAIADIDYDPLEGYVYWDDEVRATRRAYLDGSGAQLVNTNINPDG 420
Db 361 DPTDILVQVDDIRHAIADIDYDPLEGYVYWDDEVRATRRAYLDGSGAQLVNTNINPDG 420
Qy 421 IAVDWARNLYWTDGTDRIEVTRLNGTSRKILVSEDLDEPRALHPVGLMTWTWGE 480
Db 421 IAVDWARNLYWTDGTDRIEVTRLNGTSRKILVSEDLDEPRALHPVGLMTWTWGE 480
Qy 481 NPKIECANLDQERRVLVNASLGNLALDLQEGKLYMGDAKTDKIEVINVDGTKRRTL 540
Db 481 NPKIECANLDQERRVLVNASLGNLALDLQEGKLYMGDAKTDKIEVINVDGTKRRTL 540
Qy 541 LEDKLPHPFGTLTLDGFIYWTWDQRRSIRERHVKVKSARDVLIIDQLPDLMLKAVNAKV 600
Db 541 LEDKLPHPFGTLTLDGFIYWTWDQRRSIRERHVKVKSARDVLIIDQLPDLMLKAVNAKV 600
Qy 601 GTNFCADRNGCSHLCPFTPHATCGCPIGLELSDMKTCIVPEAFVFTSRAAHRISL 660
Db 601 GTNFCADRNGCSHLCPFTPHATCGCPIGLELSDMKTCIVPEAFVFTSRAAHRISL 660
Qy 661 ETNNNDVAIPLTGVKEASALDPDYSNNHIYWTDSLKTISRPFNGSGSVHEVVFGLDYP 720
Db 661 ETNNNDVAIPLTGVKEASALDPDYSNNHIYWTDSLKTISRPFNGSGSVHEVVFGLDYP 720
Qy 721 EGMAYDWGKLYWADTGTNRIEVARLDGQROVLVWRDLNPRSLALDPTKGYIYTW 780
Db 721 EGMAYDWGKLYWADTGTNRIEVARLDGQROVLVWRDLNPRSLALDPTKGYIYTW 780
Qy 781 GKGPRIVRAFMGDNCTVLDKVRANDLTIDYADQRLYWTDLTNMIESNNMLGQSRV 840
Db 781 GKGPRIVRAFMGDNCTVLDKVRANDLTIDYADQRLYWTDLTNMIESNNMLGQSRV 840
Qy 841 IADLPHFPGLTOYSYDIYWTDMNLHSIERADKTSGRNRTLIOGHLDVMDILVFHSSRQ 900
Db 841 IADLPHFPGLTOYSYDIYWTDMNLHSIERADKTSGRNRTLIOGHLDVMDILVFHSSRQ 900
Qy 901 DGLNDCMHNNGCQQLCLAIPEGHRCGCASHYTLDPSRNCSPPTTFLPSQKSAISRM 960
Db 901 DGLNDCMHNNGCQQLCLAIPEGHRCGCASHYTLDPSRNCSPPTTFLPSQKSAISRM 960
Qy 961 PDDQHSPLILPLAGLRNVAKIDYDLDKFTYWDGRQNIKRAKDDGTOPVLTSLSQG 1020
Db 961 PDDQHSPLILPLAGLRNVAKIDYDLDKFTYWDGRQNIKRAKDDGTOPVLTSLSQG 1020
Qy 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVNARGY 1080
Db 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVNARGY 1080
Qy 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWADLKRIESCD 1140
Db 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWADLKRIESCD 1140
Qy 1141 LSGANRLTLEDANIVQPLGTLKHLWIDRQOQMIERVKTTGDKRTRIQGRVAHLTG 1200
Db 1141 LSGANRLTLEDANIVQPLGTLKHLWIDRQOQMIERVKTTGDKRTRIQGRVAHLTG 1200
Qy 1201 IHAVEEVSLEBFAHPCARDNGGCSHICIAKGDTGPRCPCPVHLVLLQNLLTCGEPPTCS 1260
```

```
Db 1201 IHAVEEVSLEBFAHPCARDNGGCSHICIAKGDTGPRCPCPVHLVLLQNLLTCGEPPTCS 1260
Qy 1261 PDQFACATGEIDCTPGAWRCDFPECCDDQSDDEBECPCVCSAAQFFCARGQCVDLRLRCDGE 1320
Db 1261 PDQFACATGEIDCTPGAWRCDFPECCDDQSDDEBECPCVCSAAQFFCARGQCVDLRLRCDGE 1320
Qy 1321 ADCQDRDEADCDACILPNQPRCASGOCVLKIQCCDSPDCIDGSDLMCEITKPPSDDS 1380
Db 1321 ADCQDRDEADCDACILPNQPRCASGOCVLKIQCCDSPDCIDGSDLMCEITKPPSDDS 1380
Qy 1381 PAHSSAIGPVIGIILSLFVMGVYFVQORVVCQRYAGANGPPHYYVSGTTPHVLNFIAP 1440
Db 1381 PAHSSAIGPVIGIILSLFVMGVYFVQORVVCQRYAGANGPPHYYVSGTTPHVLNFIAP 1440
Qy 1441 GGSQHGPFPTGIACGSKMSSVSLMGGRGVPLYDRNHVHTGASSSSSSSTKATLYPPIINP 1500
Db 1441 GGSQHGPFPTGIACGSKMSSVSLMGGRGVPLYDRNHVHTGASSSSSSSTKATLYPPIINP 1500
Qy 1501 PPSPATDPSLYNMDMFYSSNIPATARPYPYIIIRGMAPPPTTSCSTDVCDSDYSASRWKAS 1560
Db 1501 PPSPATDPSLYNMDMFYSSNIPATARPYPYIIIRGMAPPPTTSCSTDVCDSDYSASRWKAS 1560
Qy 1561 KYXLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
Db 1561 KYXLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 7
US-10-182-936A-3
; Sequence 3, Application US/10182936A
; Publication No. US20040038860A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Kristina M.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Bhat, Bheem
; APPLICANT: Damagnez, Veronique
; APPLICANT: Robinson, John
; APPLICANT: Yaworsky, Paul
; TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
; FILE REFERENCE: 032796-143
; CURRENT APPLICATION NUMBER: US/10/182,936A
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: PCT/US02/15982
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-936A-3

Query Match 99.7%; Score 8715; DB 15; Length 1615;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1610; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MEAAPGPPWPLLLLLLLLLLALCCGPAPAAASPLLLFANRRDVLVDAGGVKLESTIVVS 60
Db 1 MEAAPGPPWPLLLLLLLLLLALCCGPAPAAASPLLLFANRRDVLVDAGGVKLESTIVVS 60
Qy 61 GLEDAADVDFQSKGAVYTWDSSEAIKQTYLNTGTGAQVNVVIGSLVSPDGLACDWGK 120
Db 61 GLEDAADVDFQSKGAVYTWDSSEAIKQTYLNTGTGAQVNVVIGSLVSPDGLACDWGK 120
Qy 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDPQRAIALDPAHGYMYTWDGTPRIERAG 180
```

Db 121 KLYWTDSETRNIEVANLNGTSRKVLFWQDLDPKALALDPAHYMYWTDWGETPRIERAG 180
Qy 181 MDGSTRKIIIVDSIIYWPNGLTIDLBEOQLYADAKLSFIHRANLDGSPROKVVESGLTTP 240
Db 181 MDGSTRKIIIVDSIIYWPNGLTIDLBEOQLYADAKLSFIHRANLDGSPROKVVESGLTTP 240
Qy 241 FALTLSGDTLYWTDQTRSIHACNKTGGRKEIISALYSMPMDIOVLQOEROPFFHTTCE 300
Db 241 FALTLSGDTLYWTDQTRSIHACNKTGGRKEIISALYSMPMDIOVLQOEROPFFHTTCE 300
Qy 301 EDNGCCHLCLSPSEPFYTCACPTGVQDQNGRTCKAGAEVILLARRTLRISLDT 360
Db 301 EDNGCCHLCLSPSEPFYTCACPTGVQDQNGRTCKAGAEVILLARRTLRISLDT 360
Qy 361 DFTDVLQVDDIRHAIADYDPLEGYVYWTDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
Db 361 DFTDVLQVDDIRHAIADYDPLEGYVYWTDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
Qy 421 IAVDMVARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVMGLMYWTDWGE 480
Db 421 IAVDMVARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVMGLMYWTDWGE 480
Qy 481 NPKIECANLDQERRVLVNASLWPNGLALDLQEGKLYWDAKTDKIEVINVDGFKRTL 540
Db 481 NPKIECANLDQERRVLVNASLWPNGLALDLQEGKLYWDAKTDKIEVINVDGFKRTL 540
Qy 541 LEDKLPHIFGFTLLGDFIYWTDQWRSIERVHKVKAASRDVLDIQLPDLMLKAVNAVW 600
Db 541 LEDKLPHIFGFTLLGDFIYWTDQWRSIERVHKVKAASRDVLDIQLPDLMLKAVNAVW 600
Qy 601 GTNPCADRNGGCHLCFFTPHATRCGCPGLELLSDMTKCIPEAFVFTSRAAIIHRLS 660
Db 601 GTNPCADRNGGCHLCFFTPHATRCGCPGLELLSDMTKCIPEAFVFTSRAAIIHRLS 660
Qy 661 ETNNNDVAIPITGVKEASALDFVSNHNIYWDVSLKTIISRAFMNGSSVEHVFEGLDYP 720
Db 661 ETNNNDVAIPITGVKEASALDFVSNHNIYWDVSLKTIISRAFMNGSSVEHVFEGLDYP 720
Qy 721 EGMADVMMKNIYADTCTNRIEVARLDGQFVLVWDLNPRSLALDPTKGYIYWTW 780
Db 721 EGMADVMMKNIYADTCTNRIEVARLDGQFVLVWDLNPRSLALDPTKGYIYWTW 780
Qy 781 GGPRIIVAFMDGTCNMTLVKVGANDLTIDYADQRLYWTDLTNMIESNNMLGQERV 840
Db 781 GGPRIIVAFMDGTCNMTLVKVGANDLTIDYADQRLYWTDLTNMIESNNMLGQERV 840
Qy 841 IADDLPHFPGLTOYSDIYTDWNLHSTERADKTSGRNRTLIQGLDFVMDILVFHSSRQ 900
Db 841 IADDLPHFPGLTOYSDIYTDWNLHSTERADKTSGRNRTLIQGLDFVMDILVFHSSRQ 900
Qy 901 DGLNDCHNNGCQGLCLAIPEGHRCGCASHYTLDPSSRNGSPPTTFLFSKSAISRM 960
Db 901 DGLNDCHNNGCQGLCLAIPEGHRCGCASHYTLDPSSRNGSPPTTFLFSKSAISRM 960
Qy 961 PDDQHSPLILPLHGLRNVAIDAIDYDPLDKFIYWDGQRIKRAKDDGTQPPVLTSLSQG 1020
Db 961 PDDQHSPLILPLHGLRNVAIDAIDYDPLDKFIYWDGQRIKRAKDDGTQPPVLTSLSQG 1020
Qy 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVLRGDRKPRAIIVNAERG 1080
Db 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVLRGDRKPRAIIVNAERG 1080
Qy 1081 LYFTNMQDRAAKIIRAALDGTGEREVLTFTGLIRPVVALVVDNTLGLFWADADLRIESCD 1140
Db 1081 LYFTNMQDRAAKIIRAALDGTGEREVLTFTGLIRPVVALVVDNTLGLFWADADLRIESCD 1140
Qy 1141 LSGANRLTLEDANIVQPLGLTILGKLYWIDRQOQMTERVEKTTGDKRTRIQRVAHLTG 1200
Db 1141 LSGANRLTLEDANIVQPLGLTILGKLYWIDRQOQMTERVEKTTGDKRTRIQRVAHLTG 1200
Qy 1201 IHAVEVSLEEFSAHPCARDNGGCHICIAKDGCTPRCSCPVLHVLQNLITCCEPPTCS 1260
Db 1201 IHAVEVSLEEFSAHPCARDNGGCHICIAKDGCTPRCSCPVLHVLQNLITCCEPPTCS 1260

Qy 1261 PDQFACATGEIDCIPGAWRCDFPECCDDQSBEGCPVCSAAQFPCCARQCQVDLRLRCDE 1320
Db 1261 PDQFACATGEIDCIPGAWRCDFPECCDDQSBEGCPVCSAAQFPCCARQCQVDLRLRCDE 1320
Qy 1321 ADCODRSDEADCAICLPNQFRCSAGQCVLIKQOCDSPDCIDGSDLMCEITKPPSDDS 1380
Db 1321 ADCODRSDEADCAICLPNQFRCSAGQCVLIKQOCDSPDCIDGSDLMCEITKPPSDDS 1380
Qy 1381 PAHSSAIGPVGIIILSLFVWGGVYFVQVVCQRYAGANGPPHYEYVSGTHVPLNFIAP 1440
Db 1381 PAHSSAIGPVGIIILSLFVWGGVYFVQVVCQRYAGANGPPHYEYVSGTHVPLNFIAP 1440
Qy 1441 GSGHQGPPTGIACGKSMSSVSLMGRGGVPLYDRNHVTGASSSSSSSTKATLYPPLNP 1500
Db 1441 GSGHQGPPTGIACGKSMSSVSLMGRGGVPLYDRNHVTGASSSSSSSTKATLYPPLNP 1500
Qy 1501 PPSPATPSLYNMDMEYSSNIPATARPVPIIRGMAPPPTPCSTDYCDSDYSASRWKAS 1560
Db 1501 PPSPATPSLYNMDMEYSSNIPATARPVPIIRGMAPPPTPCSTDYCDSDYSASRWKAS 1560
Qy 1561 KYILDLMSDSPPYPPPTPHSQYLSAEDSCPPSPATERSYFHLFPPPPSPCTDSS 1615
Db 1561 KYILDLMSDSPPYPPPTPHSQYLSAEDSCPPSPATERSYFHLFPPPPSPCTDSS 1615

RESULT 8

US-10-731-739-3
; Sequence 3, Application US/10731739
; Publication No. US20040176582A1
; GENERAL INFORMATION:
; APPLICANT: Carulli, John P.
; APPLICANT: Recker, Robert R.
; APPLICANT: Johnson, Mark L.
; TITLE OF INVENTION: High bone mass gene of 11q13.3
; FILE REFERENCE: 032796-013
; CURRENT APPLICATION NUMBER: US/10/731,739
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/09/544,398B
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 641
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-731-739-3

Query Match 99.7%; Score 8715; DB 16; Length 1615;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1610; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MEAAPGPPWPLLLLLLLALCCGAPAPAAASPLLLFANRRDVRVLDAGGVKLESTIWS 60
Db 1 MEAAPGPPWPLLLLLLLALCCGAPAPAAASPLLLFANRRDVRVLDAGGVKLESTIWS 60
Qy 61 GLEDAADVDFQSKGAVYTDVSEEAIKOTYINOTGAONVVISGLVSPGLACDWVGK 120
Db 61 GLEDAADVDFQSKGAVYTDVSEEAIKOTYINOTGAONVVISGLVSPGLACDWVGK 120
Qy 121 KLYWTDSETRNIEVANLNGTSRKVLFWQDLDPKALALDPAHYMYWTDWGETPRIERAG 180
Db 121 KLYWTDSETRNIEVANLNGTSRKVLFWQDLDPKALALDPAHYMYWTDWGETPRIERAG 180
Qy 181 MDGSTRKIIIVDSIIYWPNGLTIDLBEOQLYADAKLSFIHRANLDGSPROKVVESGLTTP 240
Db 181 MDGSTRKIIIVDSIIYWPNGLTIDLBEOQLYADAKLSFIHRANLDGSPROKVVESGLTTP 240

Db 181 MDGSTRKIIIVDSIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPRQKVVESGLTHP 240
Qy 241 FALTLSGDTLYWTDWQTRSIIHACNKRKTGGKKEITLSALYSMPDIQVLSQERQPPFHTRCE 300
Db 241 FALTLSGDTLYWTDWQTRSIIHACNKRKTGGKKEITLSALYSMPDIQVLSQERQPPFHTRCE 300
Qy 301 EDNGGCHLCLLSPEPPYTCACPTGVLQDNGRTCKAGAEVLLARTRDRLRLSLDTP 360
Db 301 EDNGGCHLCLLSPEPPYTCACPTGVLQDNGRTCKAGAEVLLARTRDRLRLSLDTP 360
Qy 361 DFTDIVLQVDDIRAIADIDYDPLGYVYVWTDDEVRATRRAYLDGSGAQTLVNTINPDG 420
Db 361 DFTDIVLQVDDIRAIADIDYDPLGYVYVWTDDEVRATRRAYLDGSGAQTLVNTINPDG 420
Qy 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALALHPVGMGLMTWTDGE 480
Db 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALALHPVGMGLMTWTDGE 480
Qy 481 NPKIECANLDQOERRVLVNASLGNPGLALDLQEGKLYWGDATKIEVINVDGTRKRTL 540
Db 481 NPKIECANLDQOERRVLVNASLGNPGLALDLQEGKLYWGDATKIEVINVDGTRKRTL 540
Qy 541 LEDKLPHFGEITLGDPIYWTWQRRSIEVHVKKASRDVITDQIPDLMLKAVNAKV 600
Db 541 LEDKLPHFGEITLGDPIYWTWQRRSIEVHVKKASRDVITDQIPDLMLKAVNAKV 600
Qy 601 GTNPCADRNGGCHLCPFTPHATCGCPIGLELLSDMKTCIVPEAFVFTSRAAIIHISL 660
Db 601 GTNPCADRNGGCHLCPFTPHATCGCPIGLELLSDMKTCIVPEAFVFTSRAAIIHISL 660
Qy 661 ETNNNDVAIPLTGVKEASALDFOVSNHHIYWTDVSLKTIISRAFMNGSSVEHVFEGLDYP 720
Db 661 ETNNNDVAIPLTGVKEASALDFOVSNHHIYWTDVSLKTIISRAFMNGSSVEHVFEGLDYP 720
Qy 721 EGMADVAMKGLYWADGTNRIEVARLDGQROVLVWRDLNPRSLALDPTKGIIYTW 780
Db 721 EGMADVAMKGLYWADGTNRIEVARLDGQROVLVWRDLNPRSLALDPTKGIIYTW 780
Qy 781 GKKPRIVRAFMDGNTCMTLVDKGRANDLTIDYADQRLYWTDLTNMIESNMGLQSERVV 840
Db 781 GKKPRIVRAFMDGNTCMTLVDKGRANDLTIDYADQRLYWTDLTNMIESNMGLQSERVV 840
Qy 841 IADLPHPPGLTQYSDIYIYWTMNLHLSIERADKTSGRNRTLQGHLDPMVILVFHSSRQ 900
Db 841 IADLPHPPGLTQYSDIYIYWTMNLHLSIERADKTSGRNRTLQGHLDPMVILVFHSSRQ 900
Qy 901 DGLNDCMHNNGCQCGLCIAIPGHRCCGASHYTLDPSSRNCSPPTTFLFSQKSAISRM 960
Db 901 DGLNDCMHNNGCQCGLCIAIPGHRCCGASHYTLDPSSRNCSPPTTFLFSQKSAISRM 960
Qy 961 PDDQHSPLILPLHGLNRVKAIDYDPLDKFYWDGQNKIKRAKDDGTQPPVLTSLSQG 1020
Db 961 PDDQHSPLILPLHGLNRVKAIDYDPLDKFYWDGQNKIKRAKDDGTQPPVLTSLSQG 1020
Qy 1021 NPDRQPHDLSDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
Db 1021 NPDRQPHDLSDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
Qy 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVWNTIGKLFWDADLKRIESCD 1140
Db 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVWNTIGKLFWDADLKRIESCD 1140
Qy 1141 LSGANRLTLTDANIVQPLGLTILGKHYWIDRQOQMIERVEKTTGDKRTRIQGRVAHJTG 1200
Db 1141 LSGANRLTLTDANIVQPLGLTILGKHYWIDRQOQMIERVEKTTGDKRTRIQGRVAHJTG 1200
Qy 1201 IHAVEVSLSEFSAHPCARDNGGCHSICAKGDGTPRCSPVHLVLLQNLITCGEPTCS 1260
Db 1201 IHAVEVSLSEFSAHPCARDNGGCHSICAKGDGTPRCSPVHLVLLQNLITCGEPTCS 1260
Qy 1261 PDQFACATGEIDICIPGAWRCDGPFECDDQSDDEGCPVCSAAQPPCARGQCVDLRLRCDGE 1320
Db 1261 PDQFACATGEIDICIPGAWRCDGPFECDDQSDDEGCPVCSAAQPPCARGQCVDLRLRCDGE 1320

Qy 1321 ADCODRSEADCDALCLPNQFRCASGQCVLIKQOCDSPPDCIDGSDLMCBITKPPSDDS 1380
Db 1321 ADCODRSEADCDALCLPNQFRCASGQCVLIKQOCDSPPDCIDGSDLMCBITKPPSDDS 1380
Qy 1381 PAHSSAIGPVGIIILSLFVMGVYVFCQVVVCQRYAGANGPPPHBYVSGTDPHVLNFIAP 1440
Db 1381 PAHSSAIGPVGIIILSLFVMGVYVFCQVVVCQRYAGANGPPPHBYVSGTDPHVLNFIAP 1440
Qy 1441 GGSQHGPTGTACCKSMSSVSLMGGRGVPLYDRNHVGTGASSSSSSSTKATLYPILNP 1500
Db 1441 GGSQHGPTGTACCKSMSSVSLMGGRGVPLYDRNHVGTGASSSSSSSTKATLYPILNP 1500
Qy 1501 PPSPATDPSLYNMDPFYSSNIPATAPYRPVYIIRGMAPTTTSCSTDVCDSDYSASRWKAS 1560
Db 1501 PPSPATDPSLYNMDPFYSSNIPATAPYRPVYIIRGMAPTTTSCSTDVCDSDYSASRWKAS 1560
Qy 1561 KYILDLSNDSPPYPPPTPHSQYLSAEDSCPPSPTATERSYFHLPPPPSPCTDSS 1615
Db 1561 KYILDLSNDSPPYPPPTPHSQYLSAEDSCPPSPTATERSYFHLPPPPSPCTDSS 1615

RESULT 9

US-10-374-979-4
; Sequence 4, Application US/10374979
; Publication No. US2003021979A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE OF INVENTION: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 4
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-374-979-4

Query Match 99.6%; Score 8706; DB 15; Length 1615;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1609; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy 1 MEAAPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRLLVADAGGVKLESTIVS 60
Db 1 MEAAPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRLLVADAGGVKLESTIVS 60
Qy 61 GLEDAAAVDQFSKGVYVWTDVSEEAIKOTVLTGQVAVQNVVLSGLVSPGLACDWYVK 120
Db 61 GLEDAAAVDQFSKGVYVWTDVSEEAIKOTVLTGQVAVQNVVLSGLVSPGLACDWYVK 120
Qy 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDPRAIALDPAHGYWYWTGCTPRIERAG 180
Db 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDPRAIALDPAHGYWYWTGCTPRIERAG 180
Qy 181 MDGSTRKIIIVDSIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPRQKVVESGLTHP 240
Db 181 MDGSTRKIIIVDSIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPRQKVVESGLTHP 240
Qy 241 FALTLSGDTLYWTDWQTRSIIHACNKRKTGGKKEITLSALYSMPDIQVLSQERQPPFHTRCE 300
Db 241 FALTLSGDTLYWTDWQTRSIIHACNKRKTGGKKEITLSALYSMPDIQVLSQERQPPFHTRCE 300

```
QY 301 EDNGGCSHLCLLSPSEPFYTCACPTGVOLQDNGRTCKAGAEVILLARRTDLRLSLDTP 360
Db 301 EDNGGWSHCLLSPSEPFYTCACPTGVQMDNGRTCKAGAEVILLARRTDLRLSLDTP 360
QY 361 DFTDVLQVDDIRHAIADIDYDPLEGVVYWDDEVRAIRRAYLDGSGAQLTVNTEINDPDG 420
Db 361 DFTDVLQVDDIRHAIADIDYDPLEGVVYWDDEVRAIRRAYLDGSGAQLTVNTEINDPDG 420
QY 421 TAVDVARNLYWTDGTDBRIEYTRNLNGTSRKLLVSEDLDEPRAIALHPVMGLMYWTDWGE 480
Db 421 TAVDVARNLYWTDGTDBRIEYTRNLNGTSRKLLVSEDLDEPRAIALHPVMGLMYWTDWGE 480
QY 481 NPKIECANLQGERVNLVNASLWPNGLALDLOEGLYWGAKTDKIEVINVDGTKRRTL 540
Db 481 NPKIECANLQGERVNLVNASLWPNGLALDLOEGLYWGAKTDKIEVINVDGTKRRTL 540
QY 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIERVHKVKSARDVIIDQLPDLMLGKAVNAVWV 600
Db 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIERVHKVKSARDVIIDQLPDLMLGKAVNAVWV 600
QY 601 GTNPCADRNGCGSHLCFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRISL 660
Db 601 GTNPCADRNGCGSHLCFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRISL 660
QY 661 ETNNNDVAIPLTGVKEASALDFDVSNHHIYWDVSLKTIISRAFMNGSSVEHVVERGLDYP 720
Db 661 ETNNNDVAIPLTGVKEASALDFDVSNHHIYWDVSLKTIISRAFMNGSSVEHVVERGLDYP 720
QY 721 EGMVDMWGNLYWADTGTNRIEVARLDGQFQVLRVWRDLNPRSLALDPTKGYIYWTEW 780
Db 721 EGMVDMWGNLYWADTGTNRIEVARLDGQFQVLRVWRDLNPRSLALDPTKGYIYWTEW 780
QY 781 GSKPRIVFAFWDGTCMVLTKVKGANDLTIDYADQRLYWDLTDTNMISSNMLGOERVV 840
Db 781 GSKPRIVFAFWDGTCMVLTKVKGANDLTIDYADQRLYWDLTDTNMISSNMLGOERVV 840
QY 841 IADDLPEHFGLTQYSDIYIYTDWNLHLSIERADKTSGRNRTLIQGHLDVFMVILVPHSRQ 900
Db 841 IADDLPEHFGLTQYSDIYIYTDWNLHLSIERADKTSGRNRTLIQGHLDVFMVILVPHSRQ 900
QY 901 DGLNDCMHNNGCGGOLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLFFSQSAISMI 960
Db 901 DGLNDCMHNNGCGGOLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLFFSQSAISMI 960
QY 961 PDDQHSPLIILPLHGLRNKALIDVPLDKFYIYWDGQNIKRAKDDGTQPFVLTSLSQG 1020
Db 961 PDDQHSPLIILPLHGLRNKALIDVPLDKFYIYWDGQNIKRAKDDGTQPFVLTSLSQG 1020
QY 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
Db 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
QY 1081 LYFTNMQDRAAKIERAALDGTREVEVFTTGLIRPVALVVDNTLQKLFWVDADLKRIBSCD 1140
Db 1081 LYFTNMQDRAAKIERAALDGTREVEVFTTGLIRPVALVVDNTLQKLFWVDADLKRIBSCD 1140
QY 1141 LSGANRLTLEDAIVQPLGLTILGKHLWIDRQOQMIRVEKTTGDKRTRIQGRVAHLTG 1200
Db 1141 LSGANRLTLEDAIVQPLGLTILGKHLWIDRQOQMIRVEKTTGDKRTRIQGRVAHLTG 1200
QY 1201 IHAVEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQNLLTCGEPPTCS 1260
Db 1201 IHAVEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQNLLTCGEPPTCS 1260
QY 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDSEGCPCVCSAAQFFCARGQCQVDLRLRCDGE 1320
Db 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDSEGCPCVCSAAQFFCARGQCQVDLRLRCDGE 1320
QY 1321 ADCQDRSDEADCDAILCLPNQPRCAGQCVLTKQCDGFPDCCIDGSDMLMCEITKPPSDDS 1380
Db 1321 ADCQDRSDEADCDAILCLPNQPRCAGQCVLTKQCDGFPDCCIDGSDMLMCEITKPPSDDS 1380
QY 1381 PAHSSAIGPVGIIILSLFVMGGVYFVCQVVQRYAGANGPFPHEVSGTPHPVPLNFIAP 1440
```

```
Db 1381 PAHSSAIGPVGIIILSLFVMGGVYFVCQVVQRYAGANGPFPHEVSGTPHPVPLNFIAP 1440
QY 1441 GGSQHGPPTGIACGKSMSSSVSLMGGRGVPLYDRNHVVTGASSSSSSSTKATLYPPIILNP 1500
Db 1441 GGSQHGPPTGIACGKSMSSSVSLMGGRGVPLYDRNHVVTGASSSSSSSTKATLYPPIILNP 1500
QY 1501 PPSPATDPSLNMDFYSSNIPATARPYPYIIRGMAPPTTPCSTDVCDSDYSASRWKAS 1560
Db 1501 PPSPATDPSLNMDFYSSNIPATARPYPYIIRGMAPPTTPCSTDVCDSDYSASRWKAS 1560
QY 1561 KYIYDLNLSDSPPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
Db 1561 KYIYDLNLSDSPPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 10
US-10-182-936A-4
; Sequence 4, Application US/10182936A
; Publication No. US20040038860A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Kristina M.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Bhat, Bheem
; APPLICANT: Damagnez, Veronique
; APPLICANT: Robinson, John
; APPLICANT: Yaworsky, Paul
; TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
; FILE REFERENCE: 032796-143
; CURRENT APPLICATION NUMBER: US/10/182,936A
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: PCT/US02/15982
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-936A-4

Query Match 99.6%; Score 8706; DB 15; Length 1615;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1609; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MEAAPPQPPPLLLLLLLLLLALCCPAPAAASPLLLPANRRDRLVLDAGGVKLESTIVS 60
Db 1 MEAAPPQPPPLLLLLLLLLLALCCPAPAAASPLLLPANRRDRLVLDAGGVKLESTIVS 60
QY 61 GLEDAADVDFQSKGAVYWDVSEEAIKQYLNQTAGVONVYISGLVSPDGLACDWGK 120
Db 61 GLEDAADVDFQSKGAVYWDVSEEAIKQYLNQTAGVONVYISGLVSPDGLACDWGK 120
QY 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLDQPRALDPAHGYWYWTDMGETPRIERAG 180
Db 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLDQPRALDPAHGYWYWTDMGETPRIERAG 180
QY 181 MDGSTRKIIYVDSIYWPNGLTIDLEQKLYWADAKLSFIHRANDGSRQKRVGSLTLP 240
Db 181 MDGSTRKIIYVDSIYWPNGLTIDLEQKLYWADAKLSFIHRANDGSRQKRVGSLTLP 240
QY 241 FALTLSGDTLYWTDWQTRSIIHACNKRGGKRSKLSALYSPMDIQVLSQSRQPFHTRCE 300
Db 241 FALTLSGDTLYWTDWQTRSIIHACNKRGGKRSKLSALYSPMDIQVLSQSRQPFHTRCE 300
QY 301 EDNGGCSHLCLLSPSEPFYTCACPTGVOLQDNGRTCKAGAEVILLARRTDLRLSLDTP 360
```

Db 301 EDNGGWSHLCILLSPEPFTYACPTGVQMDNGRTCKAGAEVLLARRTDLRLSLDTP 360
Qy 361 DFTDVLQVDDIRHAIADYDPLEGVYVWTDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
Db 361 DFTDVLQVDDIRHAIADYDPLEGVYVWTDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
Qy 421 IAVDWARNLWTDGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVWGLMYTWDGGE 480
Db 421 IAVDWARNLWTDGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVWGLMYTWDGGE 480
Qy 481 NPKIECANLDCQERRVLVNASLHPNGALDIOEGKLYWGAOKTDKIEVINVDGTKRTL 540
Db 481 NPKIECANLDCQERRVLVNASLHPNGALDIOEGKLYWGAOKTDKIEVINVDGTKRTL 540
Qy 541 LEDKLPHIFGFTLLGDFIYMTDMQRRSIEVHVKKASRDVLIIDQLPDLMLGLKANVAKV 600
Db 541 LEDKLPHIFGFTLLGDFIYMTDMQRRSIEVHVKKASRDVLIIDQLPDLMLGLKANVAKV 600
Qy 601 GTNFCADRNGGCSHLCFFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAHRISL 660
Db 601 GTNFCADRNGGCSHLCFFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAHRISL 660
Qy 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYVTDVSLKTIISAFMNGSSVEHVFEGLDYP 720
Db 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYVTDVSLKTIISAFMNGSSVEHVFEGLDYP 720
Qy 721 EGMVDMWGNKLYWADTGTNRIEVARLDGQFQVLVWRDLNPRSLALDPTKGYIYTW 780
Db 721 EGMVDMWGNKLYWADTGTNRIEVARLDGQFQVLVWRDLNPRSLALDPTKGYIYTW 780
Qy 781 GKGPRIVRAFMGDMGNTLVKVGANDLTIDYADQRLYTDLDTNMISSNMLGQBRVV 840
Db 781 GKGPRIVRAFMGDMGNTLVKVGANDLTIDYADQRLYTDLDTNMISSNMLGQBRVV 840
Qy 841 IADLPHPPFGLTQYSDIYVTDWNLHSTERADKTSGRNRTLIQGHLPFVMDILVFHSRQ 900
Db 841 IADLPHPPFGLTQYSDIYVTDWNLHSTERADKTSGRNRTLIQGHLPFVMDILVFHSRQ 900
Qy 901 DGLNDCMHNQCQGLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLFSQSAISRMI 960
Db 901 DGLNDCMHNQCQGLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLFSQSAISRMI 960
Qy 961 PDDQHSPLDILPLHGLRNVAIDYDPLDKFYVWDGRQNIKRAKDXDGTQPFVLTSLSGQ 1020
Db 961 PDDQHSPLDILPLHGLRNVAIDYDPLDKFYVWDGRQNIKRAKDXDGTQPFVLTSLSGQ 1020
Qy 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
Db 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
Qy 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIBSCD 1140
Db 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIBSCD 1140
Qy 1141 LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIEVERKTTGDKRTRIQGRVAHLTG 1200
Db 1141 LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIEVERKTTGDKRTRIQGRVAHLTG 1200
Qy 1201 IHAVEEVSLEFSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVQLNLTTCGEPPTCS 1260
Db 1201 IHAVEEVSLEFSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVQLNLTTCGEPPTCS 1260
Qy 1261 PDQFACATGEIDCIPGAWRCGDPPECDDQSEEGCPVCSAAQFPFCARGQCVDLRLRCDE 1320
Db 1261 PDQFACATGEIDCIPGAWRCGDPPECDDQSEEGCPVCSAAQFPFCARGQCVDLRLRCDE 1320
Qy 1321 ADCQDRSDEADCAICLPNQFRCSAGQCVLTIKQCCDFFPCIDGSDMLMCEITKPPSDDS 1380
Db 1321 ADCQDRSDEADCAICLPNQFRCSAGQCVLTIKQCCDFFPCIDGSDMLMCEITKPPSDDS 1380
Qy 1381 PAHSSAIGPVTGIIISLFPVWGGVYFVQVVCQRYAGANGFPFHEYSGTGHVPLNFIAP 1440
Db 1381 PAHSSAIGPVTGIIISLFPVWGGVYFVQVVCQRYAGANGFPFHEYSGTGHVPLNFIAP 1440

Qy 1441 GGSQHGPFTGIACGKSMSSVSLMGGRGVPLYDRNHVHTGASSSSSTKATLYPPIINP 1500
Db 1441 GGSQHGPFTGIACGKSMSSVSLMGGRGVPLYDRNHVHTGASSSSSTKATLYPPIINP 1500
Qy 1501 PPSPATDPSLYNMDMFYSSNIPATAPYRPIYIIRGMAPPPTTTCSTDVCDSDYSASRWKAS 1560
Db 1501 PPSPATDPSLYNMDMFYSSNIPATAPYRPIYIIRGMAPPPTTTCSTDVCDSDYSASRWKAS 1560
Qy 1561 KYYLIDNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
Db 1561 KYYLIDNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 11

US-10-731-739-4
; Sequence 4, Application US/10731739
; Publication No. US20040176582A1
; GENERAL INFORMATION:
; APPLICANT: Carulli, John P.
; APPLICANT: Little, Randall D.
; APPLICANT: Recker, Robert R.
; APPLICANT: Johnson, Mark L.
; TITLE OF INVENTION: High bone mass gene of 11q13.3
; FILE REFERENCE: 032796-013
; CURRENT APPLICATION NUMBER: US/10/731,739
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/09/544,398B
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 641
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-731-739-4

Query Match 99.6%; Score 8706; DB 16; Length 1615;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1609; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MEAAPPGPWPPLLLLLLLALCCGPAPAAASPLLLFANRRDRLVDAGGVKLESTIVVS 60
Db 1 MEAAPPGPWPPLLLLLLLALCCGPAPAAASPLLLFANRRDRLVDAGGVKLESTIVVS 60
Qy 61 GLEDAAAVDFQPSKGVYVWTDVSEAIKQTYLNOTGAQVNVISGLVSPDGLACDWGK 120
Db 61 GLEDAAAVDFQPSKGVYVWTDVSEAIKQTYLNOTGAQVNVISGLVSPDGLACDWGK 120
Qy 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDQPAIALDPAHGYMYWTDWETPRIERAG 180
Db 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDQPAIALDPAHGYMYWTDWETPRIERAG 180
Qy 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSRQKVGESLTHP 240
Db 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSRQKVGESLTHP 240
Qy 241 FALTLSGDTLYWTDWQTRSIIHACNKRGGKKEILSALYSPMDIOVLQSOEPPFHTTCE 300
Db 241 FALTLSGDTLYWTDWQTRSIIHACNKRGGKKEILSALYSPMDIOVLQSOEPPFHTTCE 300
Qy 301 EDNGGSHLCILLSPEPFTYACPTGVQMDNGRTCKAGAEVLLARRTDLRLSLDTP 360
Db 301 EDNGGSHLCILLSPEPFTYACPTGVQMDNGRTCKAGAEVLLARRTDLRLSLDTP 360
Qy 361 DFTDVLQVDDIRHAIADYDPLEGVYVWTDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
Db 361 DFTDVLQVDDIRHAIADYDPLEGVYVWTDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420

Db 361 DFTDIVLQVDDIRHAIADIDYDPLEGVYVWTDDEVRARRAYLDGSGAQTLYNTEINDPDG 420
Qy 421 IAVDWARNLYWTDGTGTRIEVTRNGTSRKILVSEDLDEPRALHHPVGLMYWTDWGE 480
Db 421 IAVDWARNLYWTDGTGTRIEVTRNGTSRKILVSEDLDEPRALHHPVGLMYWTDWGE 480
Qy 481 NPKIECANLDOQERRVLVNASLWPNGLALDLQEGKLYWGDATDKIEVINVDGTKRRTL 540
Db 481 NPKIECANLDOQERRVLVNASLWPNGLALDLQEGKLYWGDATDKIEVINVDGTKRRTL 540
Qy 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIERVHKVKSARDVIIDQLPDLMLKANVAVKV 600
Db 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIERVHKVKSARDVIIDQLPDLMLKANVAVKV 600
Qy 601 GTNPCADRNGGSHLCPTPHATRCGCPIGLELLSDMKTCIVPEAFVFTGSAIHRISL 660
Db 601 GTNPCADRNGGSHLCPTPHATRCGCPIGLELLSDMKTCIVPEAFVFTGSAIHRISL 660
Qy 661 ETNNNDVAIPUTGVKEASALDPDVSNHHIYMTDVSLSKTSIRAFMNGSSVEHVVEPLDYP 720
Db 661 ETNNNDVAIPUTGVKEASALDPDVSNHHIYMTDVSLSKTSIRAFMNGSSVEHVVEPLDYP 720
Qy 721 EGMVDMWGNKLYWADTGTNRIEVARLDGQFQVLRVWRDLNPRSLALDPTKGYIYWTEW 780
Db 721 EGMVDMWGNKLYWADTGTNRIEVARLDGQFQVLRVWRDLNPRSLALDPTKGYIYWTEW 780
Qy 781 GKGPRIVAFWDGTCMTLVKVGANDLTIDYADQRLYWTDLDTNMISSNMLGOERVV 840
Db 781 GKGPRIVAFWDGTCMTLVKVGANDLTIDYADQRLYWTDLDTNMISSNMLGOERVV 840
Qy 841 IADDLPHFPFGLTQYSDIYIYWTDMNLHSIERADKTSGRNRTLIQGHLPFVMDILVPHSSRQ 900
Db 841 IADDLPHFPFGLTQYSDIYIYWTDMNLHSIERADKTSGRNRTLIQGHLPFVMDILVPHSSRQ 900
Qy 901 DGLNDCMHNNGCQGLCIAIPGHRGCGCASHYTLDPSSRNCSPPTTFLFQSKAISRM 960
Db 901 DGLNDCMHNNGCQGLCIAIPGHRGCGCASHYTLDPSSRNCSPPTTFLFQSKAISRM 960
Qy 961 PDDQHSPLIILPLHGLRNKVIDPDLKFIYWDGRONIYKRAKDGTGTPVFLTSLSGQG 1020
Db 961 PDDQHSPLIILPLHGLRNKVIDPDLKFIYWDGRONIYKRAKDGTGTPVFLTSLSGQG 1020
Qy 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVNAERG 1080
Db 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVNAERG 1080
Qy 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVVALVVDNTLGLFWVDADLKRIBSCD 1140
Db 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVVALVVDNTLGLFWVDADLKRIBSCD 1140
Qy 1141 LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIERVEKTTGDKETRIOGRVAHLTG 1200
Db 1141 LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIERVEKTTGDKETRIOGRVAHLTG 1200
Qy 1201 IHAVEESLEEFSAHPCARDNGGSHICIAKGDGTGTPRCSFVHLVLLQNLITCGEPPTCS 1260
Db 1201 IHAVEESLEEFSAHPCARDNGGSHICIAKGDGTGTPRCSFVHLVLLQNLITCGEPPTCS 1260
Qy 1261 PDQFACATGEIDCIPGAWRCGDFPECDQSDQEBEGCPVCSAAQFFPCARGQCVDLRLCDGE 1320
Db 1261 PDQFACATGEIDCIPGAWRCGDFPECDQSDQEBEGCPVCSAAQFFPCARGQCVDLRLCDGE 1320
Qy 1321 ADCQDRSDEADCDALCLPNQFRASGQCVLKQCDSPDCIDGSDMLCEITPPSDDS 1380
Db 1321 ADCQDRSDEADCDALCLPNQFRASGQCVLKQCDSPDCIDGSDMLCEITPPSDDS 1380
Qy 1381 PAHSSAIGPVGIIILSLFVWGGVYFVQCVVQRYAGANGPFPHEIYSGTGPVPLNFIAP 1440
Db 1381 PAHSSAIGPVGIIILSLFVWGGVYFVQCVVQRYAGANGPFPHEIYSGTGPVPLNFIAP 1440
Qy 1441 GGSQHGPFTGIACGKSMSSVSLMGGRGVPLYDRNHVTGASSSSSTKATLYPPIINP 1500
Db 1441 GGSQHGPFTGIACGKSMSSVSLMGGRGVPLYDRNHVTGASSSSSTKATLYPPIINP 1500

Qy 1501 PPSPATDPSLNMDFYSSNIPATARPYPYRIIRGMAPPPTPCSTDVDCSDYSASRWKAS 1560
Db 1501 PPSPATDPSLNMDFYSSNIPATARPYPYRIIRGMAPPPTPCSTDVDCSDYSASRWKAS 1560
Qy 1561 KYIYDLNLSDSDPYPPPTPHSHOYLXSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
Db 1561 KYIYDLNLSDSDPYPPPTPHSHOYLXSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 12

US-10-464-368-81
; Sequence 81, Application US/10464368
; Publication No. US20040023356A1
; GENERAL INFORMATION:
; APPLICANT: Krumlauf, Robb
; APPLICANT: Ellies, Debra
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
; FILE REFERENCE: 40716-IP-017
; CURRENT APPLICATION NUMBER: US/10/464,368
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/388,970
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 81
; LENGTH: 1611
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-10-464-368-81

Query Match 99.6%; Score 8705; DB 15; Length 1611;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1611; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

Qy 1 MEAAPPPPPPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGVKLESTIVVS 60
Db 1 MEAAPPPPPPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGVKLESTIVVS 60
Qy 61 GLEDAAVVQFQSGKGVYWTDVSEBAIKQTYLNTGAAVQNVVIGSLVSPGLACDWVGK 120
Db 61 GLEDAAVVQFQSGKGVYWTDVSEBAIKQTYLNTGAAVQNVVIGSLVSPGLACDWVGK 120
Qy 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDPRAIADPAHGYMYWTDWGETPRIERAG 180
Db 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDPRAIADPAHGYMYWTDWGETPRIERAG 180
Qy 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSGFRQKVVEGSLTTP 240
Db 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSGFRQKVVEGSLTTP 240
Qy 241 FALTLSGDTLYWTDWQTRSIAHCKRTGGRKEILSALYSPMDIOVLQSOERQPFHTRCE 300
Db 241 FALTLSGDTLYWTDWQTRSIAHCKRTGGRKEILSALYSPMDIOVLQSOERQPFHTRCE 300
Qy 301 EDNGGCHLCLLSPSEFPYTCACPTGVOLQNGRTCKAGAEVLLARRTDLRLISLDTTP 360
Db 301 EDNGGCHLCLLSPSEFPYTCACPTGVOLQNGRTCKAGAEVLLARRTDLRLISLDTTP 360
Qy 361 DFTDIVLQVDDIRHAIADIDYDPLEGVYVWTDDEVRARRAYLDGSGAQTLYNTEINDPDG 420
Db 361 DFTDIVLQVDDIRHAIADIDYDPLEGVYVWTDDEVRARRAYLDGSGAQTLYNTEINDPDG 420
Qy 421 IAVDWARNLYWTDGTGTRIEVTRNGTSRKILVSEDLDEPRALHHPVGLMYWTDWGE 480
Db 421 IAVDWARNLYWTDGTGTRIEVTRNGTSRKILVSEDLDEPRALHHPVGLMYWTDWGE 480
Qy 481 NPKIECANLDOQERRVLVNASLWPNGLALDLQEGKLYWGDATDKIEVINVDGTKRRTL 540
Db 481 NPKIECANLDOQERRVLVNASLWPNGLALDLQEGKLYWGDATDKIEVINVDGTKRRTL 540
Qy 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIERVHKVKSARDVIIDQLPDLMLKANVAVKV 600
Db 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIERVHKVKSARDVIIDQLPDLMLKANVAVKV 600

Db 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIERVHKVKSARDVVIDQLPDLMLKAVNAVWV 600
Qy 601 GTNPCADRNCGGSHLCFFTPHATRCGCGIGLELLSDMTKTCIVPBAFLVFTSRAAIIHRLSL 660
Db 601 GTNPCADRNCGGSHLCFFTPHATRCGCGIGLELLSDMTKTCIVPBAFLVFTSRAAIIHRLSL 660
Qy 661 ETNNNDVAIPUTGVKEASALDFDVSNHHIYWDVSLKTIISRAFMNGSSVEHVVEFGLDYP 720
Db 661 ETNNNDVAIPUTGVKEASALDFDVSNHHIYWDVSLKTIISRAFMNGSSVEHVVEFGLDYP 720
Qy 721 EGMADVWNGKLYWADTGTNRLEIARLQGPQVILVWDLNPRSLALDPTKGYIYWTW 780
Db 721 EGMADVWNGKLYWADTGTNRLEIARLQGPQVILVWDLNPRSLALDPTKGYIYWTW 780
Qy 781 GKGPRIVRAFMDGTNCMTLVKVGANDLTIDYADQRLYTDLDTNMIESNNMLQOERVV 840
Db 781 GKGPRIVRAFMDGTNCMTLVKVGANDLTIDYADQRLYTDLDTNMIESNNMLQOERVV 840
Qy 841 IADDLPHFPFGLTQYSDIYIYTDWNLHSTERADKTSGRNRTLIQGLDFVMDILVFHSSRQ 900
Db 841 IADDLPHFPFGLTQYSDIYIYTDWNLHSTERADKTSGRNRTLIQGLDFVMDILVFHSSRQ 900
Qy 901 DGLANDCMHNGCGQLCLAIPEGHRCGCASHYTLDPSSRNCSPPTFTLLFSQKSAISRM 960
Db 901 DGLANDCMHNGCGQLCLAIPEGHRCGCASHYTLDPSSRNCSPPTFTLLFSQKSAISRM 960
Qy 961 PDDQHSPLIPLHGLRNKVAIDYDPLKFIYWDGRONIKRAKDDGTQPPFVLTSLSOGQ 1020
Db 961 PDDQHSPLIPLHGLRNKVAIDYDPLKFIYWDGRONIKRAKDDGTQPPFVLTSLSOGQ 1020
Qy 1021 NPDRQPHDLSIDIYSRTLFWCEATNTINVHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
Db 1021 NPDRQPHDLSIDIYSRTLFWCEATNTINVHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
Qy 1081 LYFTNQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKFWVDADLRKIESCD 1140
Db 1081 LYFTNQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKFWVDADLRKIESCD 1140
Qy 1141 LSGANRLTLEDANIVQPLGLTILGHLYWIDRQOQMIERVKTTGDKRTRIQQRVVAHLTG 1200
Db 1141 LSGANRLTLEDANIVQPLGLTILGHLYWIDRQOQMIERVKTTGDKRTRIQQRVVAHLTG 1200
Qy 1201 IHAVEVSLSEBSAHPCARDNGGCSHICIAKGDGTGPRCSFVHLVLLQNLLTCGEPPTCS 1260
Db 1201 IHAVEVSLSEBSAHPCARDNGGCSHICIAKGDGTGPRCSFVHLVLLQNLLTCGEPPTCS 1260
Qy 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEBGCPCVCSAAQFPCARGCQVDLRLCDGE 1320
Db 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEBGCPCVCSAAQFPCARGCQVDLRLCDGE 1320
Qy 1321 ADCQDRSEADCDALCLPNQFRCSGQCVLTKQCDSPPDCIDGSDLMCEITKPPSDDS 1380
Db 1321 ADCQDRSEADCDALCLPNQFRCSGQCVLTKQCDSPPDCIDGSDLMCEITKPPSDDS 1380
Qy 1381 PAHSSAIGFVIGIILSLFVWGGVYFVQVVCQRYAGANGFPFHEVYSGTGHVPLNFIAP 1440
Db 1381 PAHSSAIGFVIGIILSLFVWGGVYFVQVVCQRYAGANGFPFHEVYSGTGHVPLNFIAP 1440
Qy 1441 GGSQHGPTGTACGKSMSSVSLMGRRGVPLDRNHVHTGASSSSSSSTKATLYPPIINP 1500
Db 1441 GGSQHGPTGTACGKSMSSVSLMGRRGVPLDRNHVHTGASSSSSSSTKATLYPPIINP 1500
Qy 1501 PPSPATDPSLNMDFYSSNIPATAPRVPYIIRGMAPPPTTTCSTDVCDSDYSASRWKAS 1560
Db 1501 PPSPATDPSLNMDFYSSNIPATAPRVPYIIRGMAPPPTTTCSTDVCDSDYSASRWKAS 1560
Qy 1561 KYILDLNSDSPPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
Db 1561 KYILDLNSDSPPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 13

US-10-331-907-4

Sequence 4, Application US/10331907
Publication No. US20030181660A1
GENERAL INFORMATION:
APPLICANT: Todd, John A
Hess, John W
Caskey, Charles T
Cox, Roger D
Gerhold, David
Hammond, Holly
Hey, Patricia
Kawaguchi, Yoshihiko
Merriman, Tony R
Metzker, Michael L
TITLE OF INVENTION: No. US20030181660A1el LDL-Receptor
NUMBER OF SEQUENCES: 455
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon and Vanderhye
STREET: 1100 No. US20030181660A1th Glebe Road, Eighth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: US
ZIP: VA 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/331,907
FILING DATE: 31-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/402,923A
FILING DATE: 14-Feb-2001
APPLICATION NUMBER: PCT/GB98/01102
FILING DATE: 15-Apr-1998
APPLICATION NUMBER: US 60/043,553
FILING DATE: 15-Apr-1997
APPLICATION NUMBER: US 60/048,740
FILING DATE: 05-Jun-1997
ATTORNEY/AGENT INFORMATION:
NAME: B.J.Sadoff
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 620-81
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4091
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1591 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-331-907-4

Query Match 98.5%; Score 8607; DB 14; Length 1591;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 25 CPAPAAASPLLLFANRRDRLVADAGGVKLESTIVVSGLEDAADAVDFQPSKGAIVYWDVSE 84
Db 1 CPAPAAASPLLLFANRRDRLVADAGGVKLESTIVVSGLEDAADAVDFQPSKGAIVYWDVSE 60
Qy 85 EAIKQTYLNOTGAQVQNVVISGLVSPDGLACDWMGKLYWTDSETNRIEVANLNGTSRKV 144
Db 61 EAIKQTYLNOTGAQVQNVVISGLVSPDGLACDWMGKLYWTDSETNRIEVANLNGTSRKV 120
Qy 145 LFWDQLDQPAIALDPAHGYWYWDGSETPRIERAGMDGSTRKIIIVDSDIYWPNGLTIDL 204
Db 121 LFWDQLDQPAIALDPAHGYWYWDGSETPRIERAGMDGSTRKIIIVDSDIYWPNGLTIDL 180
Qy 205 EEQKLYWADAKLSFIHRANLDGSPRQKVVGSLTFPALTLSGDTLYWTDWQTSIHACN 264
Db 181 EEQKLYWADAKLSFIHRANLDGSPRQKVVGSLTFPALTLSGDTLYWTDWQTSIHACN 240

QY 265 KETGGKKEILSALYSPMDIQVLQSOERQPFHTTRCEEDNGGSHCLLSPSPFFTCAP 324
Db 241 KETGGKKEILSALYSPMDIQVLQSOERQPFHTTRCEEDNGGSHCLLSPSPFFTCAP 300
QY 325 TGVQLODNGRTCKAGAEVLLARLTDLRRLISLDPDFDIDVLQVDDIRHAIADYDPLE 384
Db 301 TGVQLODNGRTCKAGAEVLLARLTDLRRLISLDPDFDIDVLQVDDIRHAIADYDPLE 360
QY 385 GYVYTTDDVRAIRRAYLDGSAQTLVNTPEINDPGIAVDWVARNLYWTDCTDRIEYTR 444
Db 361 GYVYTTDDVRAIRRAYLDGSAQTLVNTPEINDPGIAVDWVARNLYWTDCTDRIEYTR 420
QY 445 LNGTSRKILVSDLEPRALALHPWGLMYWTDWGENPKIECANLDGOERRVLVNASLGW 504
Db 421 LNGTSRKILVSDLEPRALALHPWGLMYWTDWGENPKIECANLDGOERRVLVNASLGW 480
QY 505 PNLGALDLQEGKLYMGDAKTDKIEVINVDGTRKRTLLDKLPHIFGFTLLGDFIYWTWQ 564
Db 481 PNLGALDLQEGKLYMGDAKTDKIEVINVDGTRKRTLLDKLPHIFGFTLLGDFIYWTWQ 540
QY 565 RRSIERVHKVAKASRVDIIDLPLMGLKAVNVKVVGNPCADNRNGGCSHLCTFFTPHATR 624
Db 541 RRSIERVHKVAKASRVDIIDLPLMGLKAVNVKVVGNPCADNRNGGCSHLCTFFTPHATR 600
QY 625 CGCPGLELLSDMKTCIYPEAEFLVETSRAAIHRISLETNNDVAIPLTGKVEASALDFDV 684
Db 601 CGCPGLELLSDMKTCIYPEAEFLVETSRAAIHRISLETNNDVAIPLTGKVEASALDFDV 660
QY 685 SNHHIYTDVSLKTIISRAFPMNGSSVEHVVEFGLDYPEGMAVDMGKNLYWADTGNRIEV 744
Db 661 SNHHIYTDVSLKTIISRAFPMNGSSVEHVVEFGLDYPEGMAVDMGKNLYWADTGNRIEV 720
QY 745 ARLDQFQRLVWRDLNPRSLALPTKGYIYWTWEGGKPRIVRAFMGTGNTMLVLDKVG 804
Db 721 ARLDQFQRLVWRDLNPRSLALPTKGYIYWTWEGGKPRIVRAFMGTGNTMLVLDKVG 780
QY 805 RANDLTIDYADQRLVWTDLTNTMISSNMLGOERVVIAADLPHPPGLTOYSDIYIYWTWDM 864
Db 781 RANDLTIDYADQRLVWTDLTNTMISSNMLGOERVVIAADLPHPPGLTOYSDIYIYWTWDM 840
QY 865 LHSIERADKTSGRNRTLQGHLDYFVMDILVHSSRQDGLNDCMHNNGCGQLCLAIPEGH 924
Db 841 LHSIERADKTSGRNRTLQGHLDYFVMDILVHSSRQDGLNDCMHNNGCGQLCLAIPEGH 900
QY 925 RCGCASHYTLDPSSRNCSPPTFTLLFSQSAISRMI PDDQHSPLILPLHGLRNVKAI 984
Db 901 RCGCASHYTLDPSSRNCSPPTFTLLFSQSAISRMI PDDQHSPLILPLHGLRNVKAI 960
QY 985 DPLDKFIYVWDGRONIKRAKDDGTQPFVLTSLSQONPDROPHDLSDIYSTRTLFWTCEA 1044
Db 961 DPLDKFIYVWDGRONIKRAKDDGTQPFVLTSLSQONPDROPHDLSDIYSTRTLFWTCEA 1020
QY 1045 TINTINVRHLSGEAMGVILRGDRDKPRAI VVNAERGILYFTNMQDRAAKIERAALDGTRE 1104
Db 1021 TINTINVRHLSGEAMGVILRGDRDKPRAI VVNAERGILYFTNMQDRAAKIERAALDGTRE 1080
QY 1105 VLFTTGLIRPVALVDNTLGLKFWADADLKRIESCDLSGANRLTLEDANIVQPLGLTILG 1164
Db 1081 VLFTTGLIRPVALVDNTLGLKFWADADLKRIESCDLSGANRLTLEDANIVQPLGLTILG 1140
QY 1165 KHLVYIDRQQQMIERVEKTTGDKRTIQRVVAHLTGIIHAEVLSLEEPSAHPCARDNGGC 1224
Db 1141 KHLVYIDRQQQMIERVEKTTGDKRTIQRVVAHLTGIIHAEVLSLEEPSAHPCARDNGGC 1200
QY 1225 SHICIAKGDTGPRCSPVHLVLLQNLTCGPPTCSPDQFACATGEIDCIPGAWCRDGF 1284
Db 1201 SHICIAKGDTGPRCSPVHLVLLQNLTCGPPTCSPDQFACATGEIDCIPGAWCRDGF 1260
QY 1285 ECDQSDREGCPVCSAAGPCARGOCVRLRCDEADQCRSDEADCDALCLNPQFCA 1344
Db 1261 ECDQSDREGCPVCSAAGPCARGOCVRLRCDEADQCRSDEADCDALCLNPQFCA 1320

QY 1345 SGQCVLIKQCCDSPDCIDGSDCEITKTPSDSDSPAHSSAIGRVIIGIILSLFVWGGVY 1404
Db 1321 SGQCVLIKQCCDSPDCIDGSDCEITKTPSDSDSPAHSSAIGRVIIGIILSLFVWGGVY 1380
QY 1405 FVQCVVQORVAGANGPPPEYVSGTQHVPLNFIAPGSQHGPFTGIAAGKSMSSVSLM 1464
Db 1381 FVQCVVQORVAGANGPPPEYVSGTQHVPLNFIAPGSQHGPFTGIAAGKSMSSVSLM 1440
QY 1465 GGRGVPLVDRNHNVTGASSSSSSTKATLYPPIILNPPSPATDPSLYNMDMFYSSNIPAT 1524
Db 1441 GGRGVPLVDRNHNVTGASSSSSSTKATLYPPIILNPPSPATDPSLYNMDMFYSSNIPAT 1500
QY 1525 ARVPRPYIIRGMAVPTTTPCSTDVCDSDYASRWSKASKYLLDNLSDSDPYPPTPHSOYL 1584
Db 1501 VRPRPYIIRGMAVPTTTPCSTDVCDSDYASRWSKASKYLLDNLSDSDPYPPTPHSOYL 1560
QY 1585 SAEDSCPPSPATERSYFHLFPFPPSPCTDSS 1615
Db 1561 SAEDSCPPSPATERSYFHLFPFPPSPCTDSS 1591

RESULT 14

US-10-331-907-43

; Sequence 43, Application US/10331907

; Publication No. US20030181660A1

; GENERAL INFORMATION:

; APPLICANT: Todd, John A

; Hess, John W

; Caskey, Charles T

; Cox, Roger D

; Gerhold, David

; Hammond, Holly

; Hey, Patricia

; Kawaguchi, Yoshihiko

; Merriman, Tony R

; Metzker, Michael L

; TITLE OF INVENTION: No. US20030181660A1e1 LDL-Receptor

; NUMBER OF SEQUENCES: 455

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Nixon and Vanderhye

; STREET: 1100 No. US20030181660A1e1 Glebe Road, Eighth Floor

; CITY: Arlington

; STATE: Virginia

; COUNTRY: US

; ZIP: VA 22201-4714

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/331,907

; FILING DATE: 31-Dec-2002

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/402,923A

; FILING DATE: 14-Feb-2001

; APPLICATION NUMBER: PCT/GB98/01102

; FILING DATE: 15-APR-1998

; APPLICATION NUMBER: US 60/043,553

; FILING DATE: 15-APR-1997

; APPLICATION NUMBER: US 60/048,740

; FILING DATE: 05-JUN-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: B.J.Sadoff

; REGISTRATION NUMBER: 36,663

; REFERENCE/DOCKET NUMBER: 620-81

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 816-4091

; TELEFAX: (703) 816-4100

; INFORMATION FOR SEQ ID NO: 43:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1591 amino acids

; TYPE: amino acid

TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 43:
US-10-331-907-43

Query Match 98.5%; Score 8607; DB 14; Length 1591;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 1590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 25 CPAPAAAPLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQFSGKAVYWDVSE 84
DB 1 CPAPAAAPLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQFSGKAVYWDVSE 60

QY 85 EAIKQTYLNOTGAAQVNVVIVSGLSPDGLADWGGKLYWTDSETNRIEVANLNGTSKV 144
DB 61 EAIKQTYLNOTGAAQVNVVIVSGLSPDGLADWGGKLYWTDSETNRIEVANLNGTSKV 120

QY 145 LFQQLDQPRALDPAHGYMYTWDGETPRIERAGMDGSTRKIIIVDSIIYWPNGLTIDL 204
DB 121 LFQQLDQPRALDPAHGYMYTWDGETPRIERAGMDGSTRKIIIVDSIIYWPNGLTIDL 180

QY 205 EEOKLYWADAKLSPIHRANLDSGPKVVEGSLTHPPALTLISGDTLYWTDQTSIHACN 264
DB 181 EEOKLYWADAKLSPIHRANLDSGPKVVEGSLTHPPALTLISGDTLYWTDQTSIHACN 240

QY 265 KRTGGRKEILSALYSPMDIQVLSQERQPFHTRCEEDNGGCSHLCLLSPSEPFYTCACP 324
DB 241 KRTGGRKEILSALYSPMDIQVLSQERQPFHTRCEEDNGGCSHLCLLSPSEPFYTCACP 300

QY 325 TGVQLQNGRTCKAGAEVLLARLTDLRRISLSTPDPDITVLQVDDIRHAIADYDPLE 384
DB 301 TGVQLQNGRTCKAGAEVLLARLTDLRRISLSTPDPDITVLQVDDIRHAIADYDPLE 360

QY 385 GYVYVTDDEVAIRRAYLDGSGAOTLVNTEINDPGIADVAVARNLYWTDGTDRIEVTR 444
DB 361 GYVYVTDDEVAIRRAYLDGSGAOTLVNTEINDPGIADVAVARNLYWTDGTDRIEVTR 420

QY 445 LNGTHSRKILVSEDLDEPRAIALHPVMGLMYTWDGENPKIECANLDQERRVLNVLASLGW 504
DB 421 LNGTHSRKILVSEDLDEPRAIALHPVMGLMYTWDGENPKIECANLDQERRVLNVLASLGW 480

QY 505 PNGALDLQEGKLYWGDAKTQKIEVINVDGTYKRTLLBKLPHPFGFTLLGDFIYWDWQ 564
DB 481 PNGALDLQEGKLYWGDAKTQKIEVINVDGTYKRTLLBKLPHPFGFTLLGDFIYWDWQ 540

QY 565 RRSIERVHKVKAARDVVIDQLPDLMLGKAVNAVKNVGNPCADNRNGGCSHLCTFTPHATR 624
DB 541 RRSIERVHKVKAARDVVIDQLPDLMLGKAVNAVKNVGNPCADNRNGGCSHLCTFTPHATR 600

QY 625 CGCPIGLELLDMKTCIVPEAFVFTSRAAIIHRISLETNNNDVAIPLTGKVEASALDPDV 684
DB 601 CGCPIGLELLDMKTCIVPEAFVFTSRAAIIHRISLETNNNDVAIPLTGKVEASALDPDV 660

QY 685 SNNHIYWDVSLKTIISRAFNGSSVEHVVEGLDYPEGMAVDWNGKNLYWADGTGNRIEV 744
DB 661 SNNHIYWDVSLKTIISRAFNGSSVEHVVEGLDYPEGMAVDWNGKNLYWADGTGNRIEV 720

QY 745 ARLDGQFQVILVWRDLNPRSLDLPKGYIYWTWEGGKPRIVAFWMDGTNCMTLVKVG 804
DB 721 ARLDGQFQVILVWRDLNPRSLDLPKGYIYWTWEGGKPRIVAFWMDGTNCMTLVKVG 780

QY 805 RANDLTIDYADQRLYWTDLDTNMIESNNMLQGERVVIADLPHPFGFTQYSDIYWTWVN 864
DB 781 RANDLTIDYADQRLYWTDLDTNMIESNNMLQGERVVIADLPHPFGFTQYSDIYWTWVN 840

QY 865 LHSIERADKTSGRNRTLIQGHLDVMDILVPHSSRQDGLNDGMNNGCGQLCIAIPGGH 924
DB 841 LHSIERADKTSGRNRTLIQGHLDVMDILVPHSSRQDGLNDGMNNGCGQLCIAIPGGH 900

QY 925 RCGCASHYTLDPSSRNCSPPTTELLFSQKSAISRMIPDDQHSPLILPLHGLRNVKAIDY 984
DB 901 RCGCASHYTLDPSSRNCSPPTTELLFSQKSAISRMIPDDQHSPLILPLHGLRNVKAIDY 960

QY 985 DPLDKFIYVWDGRQNIKRAKDDGTQPFVLTSLSQONPDRQPHDLSIDIYSTRFTWTCFA 1044

DB 961 DPLDKFIYVWDGRQNIKRAKDDGTQPFVLTSLSQONPDRQPHDLSIDIYSTRFTWTCFA 1020

QY 1045 TINTINVRHLSGEANGVVLGRDRDPRRAIVNAERGYLYFTNMQDRAAKIERAALDGTRE 1104

DB 1021 TINTINVRHLSGEANGVVLGRDRDPRRAIVNAERGYLYFTNMQDRAAKIERAALDGTRE 1080

QY 1105 VLFTTGLIRPALVVDNLTGLKFWVDADLKRIESCDLSGANRLTLEDANIIVQPLGLTILG 1164

DB 1081 VLFTTGLIRPALVVDNLTGLKFWVDADLKRIESCDLSGANRLTLEDANIIVQPLGLTILG 1140

QY 1165 KHLWIDRQQQMIRVERKTTGDKETRIQGRVAHLTGIHAEVEVLEEFSAHPCARDNGCC 1224

DB 1141 KHLWIDRQQQMIRVERKTTGDKETRIQGRVAHLTGIHAEVEVLEEFSAHPCARDNGCC 1200

QY 1225 SHICIAKGDGTPRSCPVHLVLLQNLITCGBPPTCSPDQFACATGEIDCIPGAWRCDGFP 1284

DB 1201 SHICIAKGDGTPRSCPVHLVLLQNLITCGBPPTCSPDQFACATGEIDCIPGAWRCDGFP 1260

QY 1285 ECDQSDDEEGCPVCSAAQFPCCARGOCVDLRRCDEADCCDRSDEADCDALCLPNQFRCA 1344

DB 1261 ECDQSDDEEGCPVCSAAQFPCCARGOCVDLRRCDEADCCDRSDEADCDALCLPNQFRCA 1320

QY 1345 SGQCVLIKQCCDSPDCIDGSDLMCEITKPPSDSPAHSAGIIPVIGIILSLFVWGGVY 1404

DB 1321 SGQCVLIKQCCDSPDCIDGSDLMCEITKPPSDSPAHSAGIIPVIGIILSLFVWGGVY 1380

QY 1405 FVQORVVCQRYAGANGFPFHYVSGTTPHVPINFTAPGSGHQGPTGTIACGSKMSSVSLM 1464

DB 1381 FVQORVVCQRYAGANGFPFHYVSGTTPHVPINFTAPGSGHQGPTGTIACGSKMSSVSLM 1440

QY 1465 GGRGVPLVDRNHVYTGASSSSSTKATLYPILNPPSPATDPSLYNMDMFYSSNIPAT 1524

DB 1441 GGRGVPLVDRNHVYTGASSSSSTKATLYPILNPPSPATDPSLYNMDMFYSSNIPAT 1500

QY 1525 ARPYRPIIRGMAPPPTPCSTDVCDSDYSASRWKASKYIYLDLNSDSDPYPPPTPHSOYL 1584

DB 1501 ARPYRPIIRGMAPPPTPCSTDVCDSDYSASRWKASKYIYLDLNSDSDPYPPPTPHSOYL 1560

QY 1585 SAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

DB 1561 SAEDSCPPSPATERSYFHLPPPPSPCTDSS 1591

RESULT 15
US-10-331-907-29
; Sequence 29, Application US/10331907
; Publication No. US20030181660A1
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; Hess, John W
; Caskey, Charles T
; Cox, Roger D
; Gerhold, David
; Hammond, Holly
; Hey, Patricia
; Kawaguchi, Yoshihiko
; Merriman, Tony R
; Metzker, Michael L
; TITLE OF INVENTION: NO. US20030181660A1el LDL-Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and vanderhye
; STREET: 1100 No. US20030181660A1th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; ZIP: VA 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/331,907
 FILING DATE: 31-Dec-2002
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/402,923A
 FILING DATE: 14-Feb-2001
 APPLICATION NUMBER: PCT/GB98/01102
 FILING DATE: 15-APR-1998
 APPLICATION NUMBER: US 60/043,553
 FILING DATE: 15-APR-1997
 APPLICATION NUMBER: US 60/048,740
 FILING DATE: 05-JUN-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: B.J. Sadoff
 REGISTRATION NUMBER: 36,563
 REFERENCE/DOCKET NUMBER: 620-81
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)816-4091
 TELEFAX: (703)816-4100
 INFORMATION FOR SEQ ID NO: 29:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1639 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 29:
 US-10-331-907-29

Query Match 98.1%; Score 8574; DB 14; Length 1639;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1585; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY	26	PAPAAASPLLPANRRDLVDAGVKGLESTIVVSGLEDAADVFQFSGKGVYVTDVSEE	85
DB	50	PELHNSPLLPANRRDLVDAGVKGLESTIVVSGLEDAADVFQFSGKGVYVTDVSEE	109
QY	86	AIKQTYLNTQGAQVNVISGLVSPDGLACDVGKLYWTDSETNRIEVANLNGTSRKVL	145
DB	110	AIKQTYLNTQGAQVNVISGLVSPDGLACDVGKLYWTDSETNRIEVANLNGTSRKVL	169
QY	146	FWQDLQDPAIALDPAHGMVYWDWGETPRIERAGMDGSTRKIIIVDSIYVNPGLTIDLE	205
DB	170	FWQDLQDPAIALDPAHGMVYWDWGETPRIERAGMDGSTRKIIIVDSIYVNPGLTIDLE	229
QY	206	EQKLYWADAKLSFIHRANLDGSRFQKVVVGGSLTHPEALTLSGDTLYWTDWQTRSHACNK	265
DB	230	EQKLYWADAKLSFIHRANLDGSRFQKVVVGGSLTHPEALTLSGDTLYWTDWQTRSHACNK	289
QY	266	RTGGRKRKEILSALYSPMDIQVLSQERQPFPHTRCEDNGGCSHLCLLSPSFPFTCACT	325
DB	290	RTGGRKRKEILSALYSPMDIQVLSQERQPFPHTRCEDNGGCSHLCLLSPSFPFTCACT	349
QY	326	GVQLQDNGRTCKAGAEVLLARRTDLRRISLDPDFTDIIVLQVDDIRHAIADYDPLEG	385
DB	350	GVQLQDNGRTCKAGAEVLLARRTDLRRISLDPDFTDIIVLQVDDIRHAIADYDPLEG	409
QY	386	VYVYTTDDVRAIRRAYLDGSGAQTIVNTEINDPGIADVWVARNLYWTDGTRLEVTRL	445
DB	410	VYVYTTDDVRAIRRAYLDGSGAQTIVNTEINDPGIADVWVARNLYWTDGTRLEVTRL	469
QY	446	NGTSRKILVSEDLDEPRALALHPVGMGLMYWTDWGNPKIECANLDGQERRVLVNASLQWP	505
DB	470	NGTSRKILVSEDLDEPRALALHPVGMGLMYWTDWGNPKIECANLDGQERRVLVNASLQWP	529
QY	506	NGLALDLOEGKLYWGDAKTKEIVINVDGTRKRTILLEDKLPHIFGFTLLGDFIYWDWQR	565
DB	530	NGLALDLOEGKLYWGDAKTKEIVINVDGTRKRTILLEDKLPHIFGFTLLGDFIYWDWQR	589
QY	566	RSIERVHVKVASRDVIIDQLPDLMLGLKAVNVKVVGTNPNFCADNRGGCSHLCFFTPHATRC	625
DB	590	RSIERVHVKVASRDVIIDQLPDLMLGLKAVNVKVVGTNPNFCADNRGGCSHLCFFTPHATRC	649
QY	626	GCPIGLELLSDMKTCIVPEAFIVFTSRAAIIHRIISLETTNNNDVAIPLTGKESALDFDVS	685

DB	650	GCPIGLELLSDMKTCIVPEAFIVFTSRAAIIHRIISLETTNNNDVAIPLTGKESALDFDVS	709
QY	686	NNHIYWTVDVSLKTTISRAPFMNGSSVEHVVEFGDLYPEGMAVDMGKNLYWATGTNRIEVA	745
DB	710	NNHIYWTVDVSLKTTISRAPFMNGSSVEHVVEFGDLYPEGMAVDMGKNLYWATGTNRIEVA	769
QY	746	RLDGQFQVLYVRDLNPRSLALDPTKGYIYVWTEWGGKPRIVRAFMDCGTNCMTLVKVKGR	805
DB	770	RLDGQFQVLYVRDLNPRSLALDPTKGYIYVWTEWGGKPRIVRAFMDCGTNCMTLVKVKGR	829
QY	806	ANDLTIDYADQRLYWTDLTNNIESSNNMLGOERVVIAADLPHFPGLTQYSYIYVTDWNL	865
DB	830	ANDLTIDYADQRLYWTDLTNNIESSNNMLGOERVVIAADLPHFPGLTQYSYIYVTDWNL	889
QY	866	HSIERADKTSGRNRTLIGQHLDLVMDILVHFSSRDGLNDCHMNNQCGQLCLAIPGHR	925
DB	890	HSIERADKTSGRNRTLIGQHLDLVMDILVHFSSRDGLNDCHMNNQCGQLCLAIPGHR	949
QY	926	CGCASHYTLDPSSRNCSPPTTFLLPKOKSAISRMIPDOHSPDLILPLHGLRNKKAIDYD	985
DB	950	CGCASHYTLDPSSRNCSPPTTFLLPKOKSAISRMIPDOHSPDLILPLHGLRNKKAIDYD	1009
QY	986	PLDKFIYVWDGRQNIKRAKODGTQPFVLVTSLSQGNPDROPHDLSIDIYSRTLFWTCEAT	1045
DB	1010	PLDKFIYVWDGRQNIKRAKODGTQPFVLVTSLSQGNPDROPHDLSIDIYSRTLFWTCEAT	1069
QY	1046	NTINVHRLSGEAMGVLRGRDRKPRAIIVVNAERGILYFTNMQDRAAKIERAALDGTREVE	1105
DB	1070	NTINVHRLSGEAMGVLRGRDRKPRAIIVVNAERGILYFTNMQDRAAKIERAALDGTREVE	1129
QY	1106	LFTTGLLRPALVVDNTLGLKFWVDADLKRTESCDLSGANRLTLEDANIVQPLGITLIGK	1165
DB	1130	LFTTGLLRPALVVDNTLGLKFWVDADLKRTESCDLSGANRLTLEDANIVQPLGITLIGK	1189
QY	1166	HLVWIDRQQQMIERVEKTTGDKRTRIQGRVAHLTGIHAEVEVSLSEFSAHPCARDNGGCS	1225
DB	1190	HLVWIDRQQQMIERVEKTTGDKRTRIQGRVAHLTGIHAEVEVSLSEFSAHPCARDNGGCS	1249
QY	1226	HICAKGDTGPRCSPVHLVQLNLLTCGPPPTCSPOFACATGEIDICIPGAWRCDGPPE	1285
DB	1250	HICAKGDTGPRCSPVHLVQLNLLTCGPPPTCSPOFACATGEIDICIPGAWRCDGPPE	1309
QY	1286	CDDQSDDEEGCPVCSAAQPPCARGOCVDLRLRCDEADCDQSDRDEADCDALCPNQFRCAS	1345
DB	1310	CDDQSDDEEGCPVCSAAQPPCARGOCVDLRLRCDEADCDQSDRDEADCDALCPNQFRCAS	1369
QY	1346	GQCVLIRKQCCDSPDCIDGSDLMCEITKPPSDSPAHSAAIGPVIIGIILSLFVMGGVYF	1405
DB	1370	GQCVLIRKQCCDSPDCIDGSDLMCEITKPPSDSPAHSAAIGPVIIGIILSLFVMGGVYF	1429
QY	1406	VQRVVQRYVAGANGPPFHEVSGTPIVPLNFTAPGSGQHPFTGIACGKSMMSVSLMG	1465
DB	1430	VQRVVQRYVAGANGPPFHEVSGTPIVPLNFTAPGSGQHPFTGIACGKSMMSVSLMG	1489
QY	1466	GRGVPLVYDRNHVTGASSSSSSSTKATLYPILNPPSPATDPSLYNMDMYSSNIPATA	1525
DB	1490	GRGVPLVYDRNHVTGASSSSSSSTKATLYPILNPPSPATDPSLYNMDMYSSNIPATA	1549
QY	1526	RPYRPYIIRGMAPPPTPCSTDVCDSDYSASRWKASKYIYLDLNSDSDPYPPPTPHSOYLS	1585
DB	1550	RPYRPYIIRGMAPPPTPCSTDVCDSDYSASRWKASKYIYLDLNSDSDPYPPPTPHSOYLS	1609
QY	1586	AEDSCPPSPATERSYFHLFPFPPSPCTDSS	1615
DB	1610	AEDSCPPSPATERSYFHLFPFPPSPCTDSS	1639

Search completed: February 17, 2005, 01:48:28
 Job time : 188 secs